U. S. Department of Labor Employment and Training Administration Washington, D.C. 20210	CLASSIFICATION UIS
	CORRESPONDENCE SYMBOL TEUMC
	DATE June 10, 1994

DIRECTIVE : UIS INFORMATION BULLETIN NO. 15-94

TO : ALL REGIONAL ADMINISTRATORS

FROM : MARY ANN WYRSCH Mary Ann Wyrsch ERJ
Director
Unemployment Insurance Service

SUBJECT : WORKER PROFILING AND REEMPLOYMENT SERVICES TEST STATE: DEVELOPMENT OF THE MARYLAND MODEL

Attached is a copy of the above paper. As part of its role in providing technical assistance to States in implementing worker profiling and reemployment services systems as mandated by Public Law 103-152, the Department of Labor (DOL) and the Maryland Department of Employment and Economic Development (DEED) recently completed the development of an operational profiling system. Maryland was used as a "test State" to prove that the concepts contained in DOL Field Memorandum 35-94 can be developed into an operational system.

The paper focuses on implementation of a profiling mechanism, based on the use of a statistical model and detailed programming specifications. The profiling requirements document is included as an appendix to the paper. The Maryland agency intends to begin using the developed system on an operational basis during the summer of 1994.

This paper is highly recommended for dissemination to Regional Office staff and State staff engaged in the implementation of worker profiling and reemployment services systems.

Inquiries regarding this paper and DOL technical assistance may be addressed to Wayne Zajac, 202-219-5616. Questions on the Maryland profiling mechanism and implementation effort may be addressed to Carol Walter, Maryland Department of Employment and Economic Development (DEED), 410-333-5070.

Attachment

RESCISSIONS None	EXPIRATION DATE May 31, 1995

WORKER PROFILING AND REEMPLOYMENT SERVICES TEST STATE:

DEVELOPMENT OF THE MARYLAND MODEL

U.S. Department of Labor
Employment and Training Administration
Unemployment Insurance Service

June 10 1994

WORKER PROFILING AND REEMPLOYMENT SERVICES TEST STATE: DEVELOPMENT OF THE MARYLAND MODEL

I. Background

This paper expands upon UIS Information Bulletin 11-94 in describing the methods and procedures used to develop a statistical profiling model in the State of Maryland. The research contained in Unemployment Insurance Information Bulletin 4-94, "Profiling Dislocated Workers for Early Referral to Reemployment Services" was the initial basis for recommending the use of a statistical model in State Worker Profiling and Reemployment Services (WP/RS) systems. Since this research was done using national-level survey data, numerous parties expressed interest in seeing how the model would perform if applied at State and local levels using actual administrative data.

Unemployment Insurance Information Bulletin 11-94, "The Worker Profiling and Reemployment Services System: Identification Methods, Test State Analyses, and Provisions of Technical Assistance" detailed, among other things, the preliminary findings of the Test State analysis conducted by the Unemployment Insurance Service (UIS) in conjunction with the Maryland Department of Economic and Employment Development (DEED). While UIS Information Bulletin 11-94 reported the results obtained from analyzing a sample of Maryland's data, this paper traces the process used to develop the actual model that Maryland will put into statewide operation in July, 1994. Included with this chronology are descriptions of several problems and issues that were encountered in developing the Maryland model, how these were resolved, and areas of ongoing concern and potential refinement to the model.

There are two separate phases involved in using a statistical model: the developmental phase and the operational phase. The developmental phase includes all processes aimed at developing a statistically, operationally, and legally acceptable model. The operational phase includes all processes involved in using this model to identify UI claimants as part of a WP/RS system. This paper describes how the developmental phase was coordinated with the operational phase in Maryland. An interface between these two phases and the personnel responsible for them is critical to the successful design and implementation of a WP/RS system. In Maryland, a detailed set of programming specifications was the primary vehicle for creating this interface. This approach represents one way States may ensure that the statistical or screening model they use to identify dislocated workers is successfully translated into a functioning system. Maryland has elected to implement its initial WP/RS system before other States

and will operate this system for the near-term. Maryland will use this system to develop a complete WP/RS system as a "second-wave" State in conformance with the Federal WP/RS initiative.

II. Initial Procedures

A. Preliminary Research

(1) The current version of the Maryland model was the end product of several waves of analysis. The "national" model developed in UIS Information Bulletin 4-94 was used as the starting point for the Maryland Test State project. However, before it could be used in an operational environment, this model needed to be customized to fit Maryland's data, the dynamics of Maryland's labor markets, and the requirements of Maryland's data processing unit.

(2) In order to determine how the national model (in terms of the variables, NOT the actual coefficients) could be adapted to Maryland's data and labor market dynamics such that State-specific coefficient estimates could be derived, an historic sample of Maryland's data was assembled and several analyses were conducted. This historic data set covered the period from July 1, 1992 to June 30, 1993; one year of data was used in order to mitigate the effects of seasonality. Benefit exhaustion outcomes could be accurately assessed for Unemployment Insurance (UI) claimants filing in this time period. Since applying the national model to Maryland's data would require a degree of testing and experimentation, a 20% sample of data was taken for this purpose. Using a smaller data set reduced computer processing time and afforded greater flexibility in conducting the preliminary analyses. UIS Information Bulletin 11-94 summarizes the results of these analyses and describes the model that was the end product. These findings were presented to a panel of UIS actuarial staff and were favorably received.

(3) It was concluded that the 20% sample produced a satisfactory model in terms of the definitions of the variables, for example, defining "education" with categories such as "high school diploma, some college", etc. It was further concluded that the coefficients comprising the model to actually be used in statewide operation should be reestimated using the entire year's worth of historic data; this would yield a model that best depicted the historic time period. Thus, the structure of the Maryland model was developed through analysis of a 20% sample of the historic data set. However, developing the current model to be used in the operational phase entailed reestimating the model's coefficients using the entire historic data set.

(4) This progression fit very well with the administrative procedures in place in Marvland. The Maryland data processing unit requires that all programming specifications be written and

approved by all involved parties prior to the commencement of programming. Such an arrangement does not lend itself to the trial-and-error experimentation that frequently accompanies the development of a statistical model. Thus, using the smaller sample for initial experimentation and definition of the model's structure allowed for early approval of the specifications. Then, system programming and estimation of the current model could take place concurrently. The use of the specifications proved to be a critical step in making a successful transition from system design to system implementation in the Maryland Test State project.

B. Programming Specifications

(1) The programming specifications are contained in the "Maryland State Profiling Requirements Document" (PRD), which is attached to this paper as Appendix A. One of the most important challenges facing States in the implementation of the WP/RS system is developing an overall blueprint for system implementation that encompasses the wide range of processes involved, provides for successful, timely implementation, yet is not unduly complex. The PRD served such a purpose in the Maryland Test State project. The PRD describes all of the steps needed to successfully carry out the operational phase of Worker Profiling in Maryland; system programming, input, and output are described in detail. For the purposes of this paper, the entire PRD need not be paraphrased. It is critical, however, to underscore the close connections between the developmental and operational phases of Worker Profiling. These connections are essential in order to ensure that once a profiling approach is developed, it is correctly implemented. The PRD provides a clear illustration of how these connections were established in Maryland.

(2) The first connection involves the initial screens that are used to narrow the model's focus to claimants who are permanently separated and are UI eligible. Because of these screens, not all UI claimants will receive a probability score from the model. As suggested in FM 35-94, the Maryland Test provided for the exclusion of claimants who:

(a) had not received an initial UI payment;

(b) had specific recall dates;

(c) had union hiring hall agreements; or

(d) filed interstate claims;

and also for claimants who:

(e) were part of the UIS-sponsored Work Search Demonstration project in which Maryland is involved.

These screens were necessarily consistent between the developmental and operational phases of Worker Profiling, meaning that claimants fitting these criteria were screened out of all historic data sets used to develop the Maryland model. This provides consistency between the two phases of Worker Profiling. Pages 4 and 5 of the PRD describe how the initial screens were executed in the context of Maryland's data.

(3) Although consistency between the developmental and operational phases of Worker Profiling was judged to be important, there was one area in which such consistency could not be completely achieved-- missing data. The claimant data used for the developmental and operational phases of profiling came from both UI and Employment Service (ES) files. ES collects data on registrants' education and occupation. Examination of the UI data revealed that, for most of the data elements, a small portion of the observations either lacked a value for that element or contained an invalid value (i.e., tenure of less than zero years). More importantly, a sizeable portion of the UI claimants had not registered with the ES, meaning they had missing values for occupation and education. For purposes of developing the model, all observations containing missing or invalid values were excluded from the sample. While this may introduce some bias to the model, it was judged to be the best short-term solution to the problem. This issue is further addressed in Section III, Part B, "Treatment of Missing Data".

(4) This solution could not be used in the operational phase, however; this would amount to screening out permanently separated, UI-eligible claimants on the basis of missing data. Thus, the concept of "default values" was conceived. When a profiled claimant has a missing or invalid value for a data element used to calculate the model's probability, a default value is used to fill in the field(s), making all claimant observations complete. The default values were assigned with the intent of neutralizing the effects of the element(s) in question. For example, if a claimant were missing the value for tenure, the default tenure value should neither appreciably raise nor appreciably lower that claimant's probability in relation to all other claimants. The default values are shown below:

(a) Education:
High school diploma (this was the "base group" in the series of categorical variables used to model education).

(b) Tenure:
2 years (this was the median observation).

(c) SDA Industry employment change(*):
- 0% (neither increasing nor decreasing) if SIC code was invalid or missing.

- State average for each Industry Division if SIC code was valid but SDA code was invalid or missing.

(d) Occupation(*):
Structural Work (this was the "base group" in the series of categorical variables used to model occupation).

(e) SDA TUR(*):
State average TUR if SDA code was invalid or missing.

(*)- These three data elements are labor market information (LMI) indicators that are paired with codes from the UI/ES extract file. More detail on the sources and use of LMI data, including the assignment of default values, is provided in Section II, Part C, "Labor Market Information".

Again, the above approach was judged to be the best short-term solution to the problem of missing data. Maryland is instituting policies that will require all field offices, except for two offices with particularly high claim loads in proportion to ES staff, to register all UI claimants with the ES. UI and ES are collocated in the Maryland field offices. This will greatly reduce the instances of missing data in the future. In addition, management reports developed for monitoring of the WP/RS system (see pages 13-15 of the PRD) will include the frequencies of missing or invalid data among the new claims from each field office. These reports will give program directors a good idea of where data collection needs to be improved, further reducing the instances of missing data.

(5) Another connection between the developmental and operational phases of Worker Profiling involves the conversion of the data elements as they appear on the UI/ES extract file into actual variables to be used in both developing and implementing the statistical model. An example of such a conversion would be using the data element "years of education" to form the series of categorical variables representing "high school diploma", "some college", etc. Discerning how the data elements are best incorporated into a statistical model typically involves a degree of trial-and-error experimentation involving similar conversions. In Maryland, this experimentation was done using a 20% sample of the historic data set and produced the variable definitions used in the current version of the model. (See UIS Information

Bulletin 11-94 for more detail on this subject). The program depicted on pages 6-8 of the PRD shows how the data elements on the UI/ES extract file were converted to match these definitions in both the developmental and operational phases of the Maryland project.

(6) The final connection between the developmental and operational phases involves the actual calculation of the probability values predicted by the statistical model. Successful calculations required a common understanding of the step-by-step mechanics and operations underlying the logistic regression equation. Such an understanding was reached through communication between statistical and programming personnel and through the development of "pseudo-code", a line-by-line description of the logic used to derive the model's probability value. This pseudo-code, found on pages 9-11 of the PRD, was readily translated into actual code.

(7) As items 1-6 above illustrate, an appreciable amount of research and planning went into the development and operation of the Maryland profiling model. As part of the PRD (see pages 1-2), the procedures and deadlines to be observed were agreed upon by all personnel involved with the Maryland Test State project; all of these deadlines were subsequently met. At this point, system programming could begin and the current version of the model could be estimated.

Values of some fields used in the operational phase were not available at the outset of system programming. For example, coefficients were not available because the current version of the model had not yet been estimated. Provisions for such fields, shown below, were made in the PRD. None of these values were necessary for the early stages of system programming. Actual values were provided at a later date (see PRD, pages 21-22):

(a) Table of final values for all coefficients used in the model. Section V, "Current Version of the Maryland Model" describes the estimation process.

(b) Table of industry percent employment changes by SDA reflecting most current information available.

(c) Table of total unemployment rates by SDA reflecting most current information available.

(d) Table of final default values.

Although the model structure was agreed upon and recorded in the PRD, it was still necessary to go through all of the steps involved in developing a statistical model. Beginning with Section III "Historic Data", this paper traces these steps.

C. Labor Market Information

First, however, it is important to describe the key role that labor market information (LMI) plays in the WP/RS system. Both the national analysis (UIS Information Bulletin 4-94) and the initial Maryland analysis (UIS Information Bulletin 11-94) included three pieces of labor market information (LMI) that proved to be useful in identifying UI claimants likely to exhaust basic benefits: industry employment change, occupation employment change, and local unemployment rate. These LMI indicators are all used in some capacity in the current version of the Maryland model. Before final estimation of the model, and as part of the PRD requirements, it was necessary to decide on the source and the specific formats (e.g., time period, level of aggregation) for each of these data elements in both the developmental and operational phases. This section describes potential sources of labor market information and how these sources were used in Maryland.

1. Sources of Labor Market Information

State LMI Units: States have as potential data sources the LMI units that provide labor market information to the Bureau of Labor Statistics (BLS), to the SESA, and to the general public. These units publish reports on employment trends within industries, occupations, and sub-state areas on a periodic basis and may possess some of the data elements needed to establish an initial WP/RS system. In addition, State LMI units may possess a range of data elements beyond those used in the national or Maryland analyses and may also have personnel who would be well-suited to assist in developing a statistical model. SESA program units are encouraged to use LMI units as sources of knowledge and, perhaps, of data in the development of their WP/RS systems.

Bureau of Labor Statistics: Beginning in September, 1994, SESAs also will have the option of using labor market information provided by BLS, in conjunction with UIS, in support of the WP/RS initiative. These data will be readily available to States in a fixed format and will support State models derived from the DOL model. This arrangement will offer States the additional advantage of using data that have undergone BLS quality control procedures. The BLS/UIS arrangement will provide for distribution of the following data elements:

(1) Quarterly employment changes within industries, aggregated at sub-state levels.

(2) Annual employment changes within occupations, aggregated at the State level.

(3) Four-quarter moving average unemployment rates, aggregated at sub-state levels.

These data will be taken from the ES-202 data, from the source data for BLS' forthcoming LASER system, and from BLS' Local Area Unemployment Statistics (LAUS) data. The data are initially provided to BLS by State LMI units and undergo additional quality control procedures. The data will be available in fixed formats on an annual basis so that State models can be updated to reflect current employment trends. In addition, the common distribution of these data will facilitate the transfer of methods and ideas between States working to develop and improve their WP/RS systems. Thus, for many SESAs, the BLS data may assist expeditious development and implementation of a WP/RS system. The following sections summarize how labor market information was incorporated into the Maryland Test State project.

2. Industry Employment Change

(a) In Maryland, the BLS ES-202 data were used to derive sub-state indicators of industry employment change. The levels of aggregation used were (SIC) Industry Divisions and Service Delivery Areas (SDAs). Thus, for each of Maryland's 12 SDAs, a local measure of the recent employment change for each Industry Division was derived. The specifics of how these data were incorporated into development of the model can be found in Section IV.

(b) It was felt that, in Maryland, the BLS data represented the best option because of time constraints that were involved. With such an aggressive schedule to meet (two weeks were allotted for model development), it was most convenient for the UIS technical assistance staff to receive the data from BLS in an agreed-upon format. Once a nationwide delivery system for the BLS data is in place in September, all States will have the option of receiving these data from UIS on a computer disk.

3. Occupation Employment Change

(a) In Maryland, Dictionary of Occupational Titles (DOT) codes are used by the ES to classify registrants' former occupations. Since BLS primarily uses the Occupational Employment Statistics (OES) coding scheme, occupational employment data were not as readily available as industry data. In the nationwide delivery system, BLS will utilize a "crosswalk" between the OES and DOT coding schemes to provide States that collect DOT codes with the appropriate data on State-level occupational employment changes.

(b) Current data on occupational employment changes were not available from the Maryland LMI office either. It is anticipated that many States will initially be in a similar situation; occupational data are difficult to collect and maintain with a reasonable degree of accuracy. As a temporary measure, UIS staff

elected to use 1-digit DOT codes to create a set of nine categorical variables in the Maryland model. These enter the model in the same way as the education variables.

However, defining occupation categorically does not have the same logical value as defining education categorically. One reason is that statistical methods require that, for each set of categorical variables, one group be left out of the equation. This "base group" should be the average or typical group to which all other groups can be most meaningfully compared. For education, the rather obvious choice is the "high school diploma" group. However, for occupation there is no obvious choice among the 1-digit DOT groupings. Structural Work (construction) was selected because it was well-represented (15% of the overall population) and had a benefit exhaustion rate (51%) very close to that of the overall population (52%). Maryland DEED staff plan to test occupational employment data for inclusion in subsequent updates of the model as such data become available.

4. Unemployment Rate

(a) The Maryland model used unemployment rate data supplied by the Maryland Office of Labor Market Analysis and Information. These data measured total unemployment rate (TUR) and were initially aggregated at the county level. The specifics of how these data were incorporated into development of the model can be found in Section IV.

(b) These data were used because they were immediately available on disk in the Maryland DEED office. Although the BLS LAUS data could have been provided in a similar format, it was most convenient for UIS staff to use the in-house data for development of the Maryland model.

5. Default Values

(a) As mentioned in Section II, Part B, "Programming Specifications", each data element used in the model was assigned a default value for use in the operational phase when missing data were encountered. Thus, default values had to be assigned for the three LMI data elements.

(b) This proved to be one of the more problematic areas of the Test State initiative. The default values for the LMI elements are as follows:

(1) SDA Industry employment change was assigned a value of 0% if the SIC code was either invalid or missing. Without any information concerning a claimant's industry, this was judged to be the most neutral value. However, if a valid SIC code was given, but a missing or invalid FIPS code prevented identification of a

claimant's area of residence, the employment change variable was set equal to the State average for the claimant's Industry Division.

(2) Occupation was assigned the value of the base group, Structural Work. This is probably the weakest point of the initial Maryland model and is attributable to the categorical specification of occupation. Using "high school diploma" as the default value for education has a degree of intuitive value in terms of neutrality; using Structural Work as the default value for occupation has no such value. However, since the exhaustion rate for Structural Work was close to that of the overall population, this was as close to a neutral value as could be achieved.

(3) SDA TUR was assigned the State average TUR if the SDA code was invalid or missing.

III. Historic data

A. Description of Sample

(1) An entire year of historic claims data was used to develop the current version of the Maryland model. In Maryland, the historic records were stored in the Maryland Automated Benefits System (MABS) database. MABS contains records of all initial claims that were filed during the period used for the analysis (July 1, 1992 - June 30, 1993). However, MABS does not contain all the data elements used for the analysis. Occupation and education data for individuals are collected by ES and stored in the Job Service Applicant File database. In order to include occupation and education in the analysis, the individual MABS records had to be matched by Social Security Number with records from the Job Service applicant file. Records of claimants who had not registered with ES contained blank fields for occupation and education.

(2) Using SAS, a statistical software package available on the Maryland mainframe, the historic records were extracted from MABS and the Job Service database and combined into a single extract file. Excluding interstate claims, there were approximately 193,000 initial claims filed during the historic time period. Short programs were written in SAS to execute the remaining initial screens as specified in the PRD, and to include in the UI/ES extract file only the data elements that would be needed to develop the statistical model.

(3) The sample of 193,000 observations was reduced to approximately 90,000 through the execution of the remaining initial screens (first payment, recall, union hall, Work Search Demo participation). These 90,000 observations represent the

universe of claimants from the historic period who WOULD BE profiled by the statistical model IF they were current claimants. The data elements included for each observation in the extract file were as follows:

Data Element	Field Position	Field Length
SSN	01-09	9
Local Office Code	10-11	2
FIPS Code	12-16	5
Weekly Benefit Amount	17-22	6.2
Actual Benefit Amount	23-30	8.2
SIC Code of most recent base-period employer	31-36	6
Highest Grade Completed	37-38	2
DOT Code(*)	39-47	9
Months of Tenure with most recent base-period employer	48-50	3

(*)- when the data were later read into SPSS, only the first three positions of this code were used.

B. Treatment of Missing Data

(1) As mentioned, records of claimants who had not registered with ES were missing data for occupation and education. In addition, most of the data elements were missing or invalid for at least some portion of the observations. In order to estimate a statistical model, full data is needed for all observations. Thus, it was decided that the best short-term solution was to exclude all records containing missing data. Such exclusions have the potential to introduce bias to the model.

(2) If the excluded observations are selected randomly, no bias should result. However, for occupation and education, the two elements most frequently missing from the observations, this is not the case. Exclusions based on missing values for these two elements are not random; all claimants who did NOT register with ES are excluded. Thus, the factor(s) that determine ES registration in Maryland represent the area(s) of the model's bias. In Maryland, the proportion of UI claimants who register with ES is chiefly a function of resource levels, staffing, and administrative procedures within field offices; individual

claimant characteristics are NOT the basis for ES registration. Thus, the model's bias is primarily geographic, meaning that field offices with extremely low ES registration rates are underrepresented in the sample, and vice-versa. Beginning in July, 1994, Maryland will institute policies requiring a 100% ES registration rate in all but two field offices. This should eventually eliminate this bias in the model and will be reflected in subsequent updates of the model.

(3) Still using SAS on the mainframe, the sample of 90,000 "eligible" records was reduced to approximately 48,000 by excluding records of claimants who had not registered with ES or had missing values for tenure. The large majority of these exclusions were of non-ES registrants. Analysis of the means (averages) and frequencies of variables in the two samples showed that the main differences between the two samples were field office-based; SDAs containing field offices with extremely low ES registration rates were underrepresented, and vice-versa.

At this point, the 48,000-record extract file was downloaded from the mainframe onto two 3.5" disks and loaded onto a PC containing SPSS, another statistical software package which Maryland had. This was done because this particular version of SPSS was supplemented by an advanced statistics module that could execute the "logistic regression" procedure. Other statistical packages such as SAS, LIMDEP and NCSS are also capable of supporting this procedure. After downloading, more analyses of means and frequencies were conducted, resulting in additional exclusions based on missing values for all other data elements. Also, in keeping with the PRD, observations with certain invalid codes or with tenure values in excess of 60 years were excluded as well. This reduced the sample to a final total of 43,197 observations containing valid values for all data elements.

IV. Data Transformations

The data elements as they appeared on the UI/ES extract file could not immediately be used to conduct the estimation of the statistical model. Most of the elements needed to be converted in order to fit the specifications of the PRD. Further, two were used as "keys" for attaching LMI indicators, which were contained in separate files, to the proper records in the extract file. These procedures are described below.

A. Conversions

(1) Since the variable formats had already been determined and were incorporated into the PRD, no additional experimentation was necessary. The exact specifications used to convert the data elements on the extract file into the formats needed for the model are shown on pages 6-8 of the PRD. The conversion process involved:

(a) Grouping the values from the "Highest Grade Completed" field into the five education categories.

(b) Dividing "Tenure in Months" by 12, then truncating (rounding downward) to the last full year completed to obtain "Tenure in Years".

(c) Grouping the values for 6-digit SIC codes into their proper Industry Division (1-digit) categories.

(d) Grouping the values for 3-digit DOT codes into their proper one-digit categories.

(e) Grouping the values for FIPS codes into their proper SDA categories.

(f) Using the "Weekly Benefit Amount" and "Actual Amount Paid" fields to discern whether or not each claimant exhausted basic benefits. If the calculation (ACTUAL AMOUNT)/(26 X WEEKLY AMOUNT) produced a value greater than or equal to 1, the claimant was deemed to have exhausted basic benefits.

(2) Once these conversions were completed, the data were in the formats needed to conduct the final estimation of the statistical model. However, the claimant records were not yet complete. Two pieces of labor market information-- SDA industry employment change and SDA unemployment rate-- are used in the Maryland model and needed to be attached to each claimant record. Section B below describes this procedure.

B. <u>LMI: Industry Employment Change</u>

(1) Information on the industry employment changes within each SDA was obtained from the Bureau of Labor Statistics' ES 202 data. Since the time period depicted by the historic file ranged from July 1, 1992 through June 30, 1993, the employment information for second quarter 1992 (92.2) and second quarter 1993 (93.2) was used in the developmental phase.

(2) A file containing this information was downloaded onto a 3.5" disk in ASCII text format. This text file was read into Lotus 1-2-3 as a spreadsheet and several transformations were done. The file contained monthly employment figures for the two quarters noted above for SIC Industry Divisions within each Maryland SDA. (These were the levels of SIC and sub-state aggregation selected for initial implementation. Future research and experience may suggest different levels of aggregation). The monthly employment figures were used to calculate quarterly average employment figures for each SDA (all Industry Divisions) and for each Industry Division within the SDA.

(3) The percent employment change between second quarter 1992 and second quarter 1993 was then calculated for each Industry Division. This percent change is the actual variable used by the model to incorporate the impact of claimants' former industry on benefit exhaustion. However, preliminary testing (prior to development of the PRD) had revealed some problems with this approach. Employment in certain Industry Divisions, primarily agriculture, mining, and miscellaneous, could be very low in any given SDA. As a result, the percent employment change was exaggerated with respect to the other Industry Divisions. For example, if average mining employment declined from 20 to 10 in a given SDA, then the variable used in estimating the model would be -50%. An examination of the data showed such observations to be extreme outliers, observations that can severely damage the accuracy of a statistical model. It was decided that for such extreme cases, the percent employment change should be weighted by the ratio of employment in that Industry Division to employment in the SDA. Industry Divisions comprising less than 3 percent of SDA employment were deemed "extreme" and the percent employment changes were weighted in this manner.

(4) The converted employment change data were written out from Lotus into another ASCII text file, which could then be read into SPSS. This layout of this file was as follows:

Data Element	Field Position	Field Length
SDA code	06-08	3
SIC Industry Division code	09-10	2
Total employment (92.2)	11-16	6
Total employment (93.2)	18-23	6
Percent employment change (92.2 - 93.2)	25-30	6
Ratio of Industry Division employment to SDA employment	33-38	6
Weighted percent employment change	41-46	6

This file was read into SPSS and was then "matched" with the converted UI/ES file. The SIC Industry Division codes and SDA codes in the claimant file were used as the keys for this match. For each claimant observation in the UI/ES file, this added all of the information from the above LMI file to the claimant record.

C. LMI: SDA Unemployment Rate

(1) The information on unemployment rates within each SDA was obtained from the Maryland Office of Labor Market Analysis. This information was obtained for second quarter 1992 through second quarter 1993 in order to match the historic time period.

(2) A file containing the unemployment rates was available in a Lotus spreadsheet on a 3.5" disk. This file was loaded in and several transformations were done. The file contained monthly labor force, employment, and unemployment figures for the entire historic time period for each of Maryland's 23 counties. The counties were grouped into Maryland's 12 SDAs, and total unemployment rates corresponding to the historic time period were calculated for each SDA and for the State as a whole. Future research and experience may suggest different measures of local unemployment (e.g., insured unemployment rate, moving averages, etc.)

(3) The converted unemployment rate data were written out from Lotus into an ASCII text file, which could then be read into SPSS. This layout of this file was as follows:

Data Element	Field Position	Field Length
SDA code	05-06	2
SDA Unemployment Rate	08-11	4.1

This file was read into SPSS and was then "matched" with the converted UI/ES file. The SDA codes in the claimant file were used as the keys for this match. For each claimant observation in the UI/ES file, this added all of the information from the above LMI file to the claimant record.

At this point, all necessary data had been added to the UI/ES file within SPSS and were in the formats needed to conduct the final estimation of the statistical model. Section V below briefly describes the procedure used to conduct this estimation and how the output was integrated into the operational phase of the Maryland project.

V. Current Version of the Maryland Model

A. Estimating the Coefficients

(1) Once the transformations described in Section IV were done, the claimant file contained complete records of 43,197 permanently separated, UI-eligible claimants. The coefficients of the model were estimated using this entire sample. The statistical procedure used to conduct this estimation is known as "logistic regression" or "logit". This procedure has been used

by UIS staff in developing both the national model and the Maryland model. Logistic regression has certain statistical properties that make it particularly useful in predicting the probability of UI benefit exhaustion and has proven effective in making such predictions. Other statistical methods may also be useful in identifying claimants likely to exhaust basic benefits.

(2) Typically, when developing a statistical model, several versions are estimated and numerous tests are conducted in order to discern the optimal structure of the model. However, such experimentation was not necessary in this case. The Maryland model's structure had already been determined based on the analysis of a smaller sample of historic data (see UIS Information Bulletin 11-94). All that was necessary in this case was to write SPSS code that would execute the logistic regression procedure using the entire data set. This code, along with all SPSS code used in developing the current version of the Maryland model, is shown in Appendix B, "SPSS Code Used to Develop the Maryland Model".

B. <u>Integrating the Coefficients into the Operational Phase</u>

(1) The output of the estimation procedure is shown in Appendix C, "SPSS Output from Estimation of the Maryland Model". The estimated coefficients are shown in the first column of the table on page 3 of the appendix. These coefficients correspond (not in exact order) to the "coefficient card" fields specified in the pseudo-code on pages 9-11 of the PRD.

(2) One apparent inconsistency is that there are 18 coefficient card fields specified in the pseudo-code but only 16 estimated coefficients. This is because the values of coefficient cards 2 and 15 correspond to "high school diploma", and "Structural Work", respectively. As mentioned in Section II of this paper and in UIS Information Bulletin 11-94, these are the "base groups" for education and occupation. Statistical methods require that the coefficients of these variables be set to zero. SPSS accounts for this internally and therefore does not include these null coefficients in the output listing.

(3) The values of the 16 estimated coefficients and the 2 zero values were manually entered into the Profiling Program to be used in the operational phase. At this point, the model had been fully transferred from the developmental phase to the operational phase and system testing could begin. This testing included actual production runs and generation of all system outputs: the error report, the ranking report, and 3 management reports. Examples of these outputs are shown on pages 24-31 of the PRD.

ATTACHMENTS:

- **LIST OF REFERENCES**

- **APPENDIX A: "MARYLAND STATE PROFILING REQUIREMENTS DOCUMENT" (PRD)**

- **APPENDIX B: "SPSS CODE USED TO DEVELOP THE MARYLAND MODEL"**

- **APPENDIX C: "SPSS OUTPUT FROM ESTIMATION OF THE MARYLAND MODEL"**

REFERENCES

Profiling Legislation

A) Public Law 103 - 6, Section 4, Profiling of New Claimants
B) Public Law 103-152, Section 4, Worker Profiling

Legislative Interpretation

A) The Unemployment Compensation Amendments of 1993 (Public Law 103-152) - Provisions Affecting the Federal-State Unemployment Compensation Program. Unemployment Insurance Program Letter 13-94, January 28, 1994.

B) Draft Language - Failure to Participate in Reemployment Services. Unemployment Insurance Program Letter 13-94 Change 1, April 15, 1994.

Implementation of a Worker Profiling and Reemployment Service System

A) Profiling of Unemployment Insurance (UI) Claimants, Unemployment Insurance Program Letter 45-93, ...

B) Implementation of a System of Profiling Unemployment Insurance (UI) Claimants and Providing Them with Reemployment Services. Field Memorandum 35-94, March 22, 1994.

C) Supplement No. 1 -- Questions and Answers Supplementing Field Memorandum No. 35-94, Implementation of a System of Profiling Unemployment Insurance (UI) Claimants and Providing Reemployment Services. Field Memorandum No. 35-94 Change 1, May 2, 1994.

Profiling Mechanisms

A) Profiling Modeling Paper - Profiling Dislocated Workers for Early Referral to Reemployment Services, by Kelleen Worden. UIS Information Bulletin 4-94, January 5, 1994.

B) The Worker Profiling and Reemployment Assistance System: Identification Methods, Test State Analyses and Provision of Technical Assistance. UIS Information Bulletin No. 11-94, May 4, 1994.

Research on Profiling and Reemployment Services

Department of Labor (DOL) Report, Reemployment Services: A Review of Their Effectiveness. UIS Informational Bulletin, May 23, 1994.

APPENDIX A: MARYLAND STATE PROFILING REQUIREMENTS DOCUMENT (PRD)

MARYLAND DEPARTMENT

OF

ECONOMIC AND EMPLOYMENT DEVELOPMENT

Division of Employment and Training

MARYLAND STATE

PROFILING REQUIREMENTS DOCUMENT

Prepared by: **Linda Cartrette, OIT**
John Mingee, DOL/TAT
Kurt Schlauch, DOL/TAT
Catherine Fratus, DOL/TAT
Helen Pan, DOL/TAT
Michael Morrissey, DOL/TAT

FINAL - May 25, 1994

MARYLAND PROFILING EFFORT

Table of Contents

MARYLAND PROFILING EFFORT

PURPOSE. The purpose of this project is to implement a profiling system which ranks unemployment insurance claimants, by each individual's probability of exhausting benefits, in order to target reemployment services to those most in need. This effort includes only those items related to the on-going profiling process and does not include the efforts associated with the historical analysis to define the coefficients.

APPROACH. To accomplish this effort, there will be five basic steps: definition of requirements, development of programs and processes, testing of programs and processes, preparation of programs and documentation for production, and definition of post implementation requirements.

Requirements. Define the specifications for the creation of the profiling process.

- Initial Draft	04/29/94
- Revised Draft	05/02/94
- Final	05/04/94

Develop Programs. Programmatically create the profiling process programs needed.

- Receive Draft Coeff/Rates/Defaults	05/03/94
- Create Control Cards	05/04/94
- Create JCL stream	05/09/94
- Extract Program	05/09/94
- Conversion Program	05/11/94
- Profiling Program	05/13/94
- Ranking Report	05/16/94
- Mgt Rpt - Number of Profiled Claimants Report	05/17/94
- Mgt Rpt - Invalid Claimant Data Trends Report	05/18/94
- Mgt Rpt - Benchmark Probability Report	05/19/94

Test. Validate that the programs meet the functional requirements through testing. Testing will be performed on two types of data: full UI and JS test files, and a created set of test files. The created files will test all data possibilities and force all functions of the program to be performed.

- Develop Test Job Streams (JCL)	05/12/94
- Develop Test Directives and Cases	05/13/94
- Develop Test Data	05/18/94
- Receive/Load Actual Coeff/Rates/Defaults	05/19/94
- Exercise Tests Against UI and JS Test Data	05/20/94
- Exercise Tests Against Created Data	05/25/94

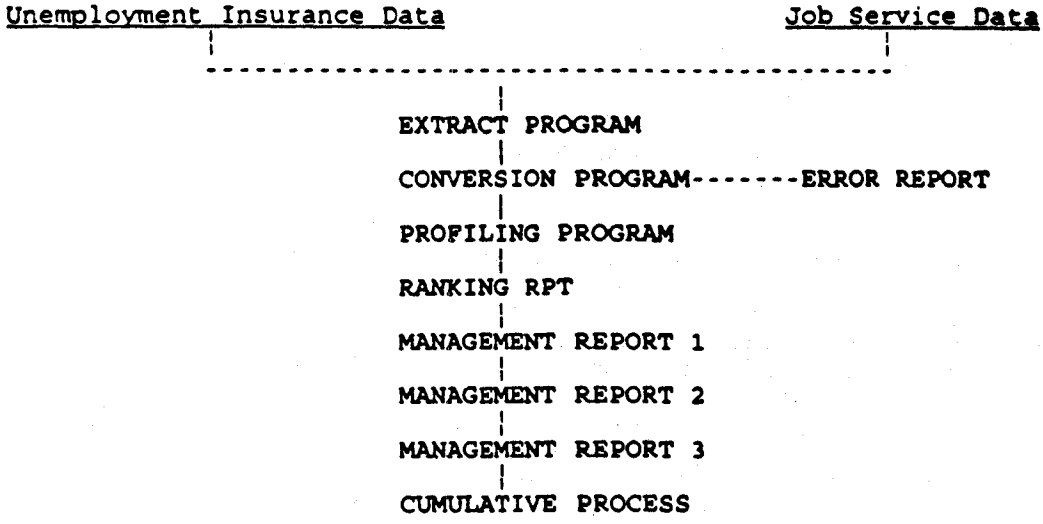
Prepare/Install. Before implementation of the profiling process in a Maryland production environment, documentation must be prepared and production approval must be obtained.

- Request Production File Names from Annapolis Data Center (ADC)	05/06/94
- Request Cylinder Space for Files From ADC	05/13/94
- Create Production Job Streams (JCL)	05/20/94
- Create Production Data Sets	05/20/94
- Create Control Procedures for ADC and Users	05/25/94
- Request Production Approval for JCL from ADC	05/26/94

Post Implementation - Once the profiling process has been installed in production, additional processes must be reviewed for development and implementation. This section highlights the processes identified during the profiling analysis and defines due dates for the initial identification and preliminary development* of requirements to support these processes.

- Define requirements for the automated update of the profiling date control card (PROFDTE). 05/05/94
- Define requirements for a downloadable file from the Profile Extract file for each Local Office. 05/05/94
- Define requirements for address labels to support the associated Ranking Report. 05/05/94
- Define requirements for the automation of the service delivery area (SDA) sequential file LMIDATA which houses unemployment rate and the SDA sequential file which house the Industry percent of change. This data will be derived from the National Office or Maryland State Office labor market information and loaded into the appropriate files. 05/06/94
- Define requirements to enhance the identification of the Service Delivery Area (SDA) code from the county (FIPS) code. The current translation converts in-state counties to SDA codes. This enhancement would provide a further translation from the local office number for the out-of-state claimants. 05/06/94
- Define requirements for the feed back process to track that the profiled claimants are reporting to job services and participating in required services. 05/10/94
- Define overall objective of the JTPA reporting process, and how that interfaces with and affects the profiling process. 05/10/94

PROFILING PROCESS FLOW CHART



EXTRACT PROGRAM: This process takes the data from the existing unemployment insurance data files and the job services data files and consolidates the data into one extract file in sequential file format.

Input File Desc: Unemployment Insurance data.
Input File 1: EUZD.TEST.BENFTMTR.EUZ960F1 (Primary) 000-217
EUZD.TEST.BENFTMTR.EUZ960F2 (Secondary)
EUZD.TEST.BENFTMTR.EUZ960F3 (Primary) 218-999
Input Format: VSAM file format
Sort Criteria: Already sorted by SSN

Input File Desc: Job Services data.
Input File 2: EMNV.ENDS.APPDATA.CASAC-AR or
EMND.ENDT.VSCLST.APP (test file)
Input Format: VSAM file format
Sort Criteria: Already sorted by SSN

Input File Desc: Date Range Control Card
Input File 3: EMNP.ENDS.CTLCDS(PROFDTE)
Input Format: Partitioned Data Set file format
Sort Criteria: Not Applicable

Output File Desc: Profiling Extract File
Output File 1: EMNP.ENDS.UIJS.DATA
Output Format: Sequential file format (permanent)
Sort Criteria: Sort by SSN

EXCLUSION CRITERIA: This section defines the input record fields selected, the validation criteria, and the status of the record for inclusion or exclusion based on that validation.

Input Field	Value	Exclusion Criteria/Reason
CPD-LOCAL-OFFICE	50-56,58,59 91,94,95,97	Exclude claimants with these values. They do not represent local office values and records are not applicable to the profiling effort. Office number 50 and 56 meet the interstate exclusion criteria.
CPD-EMPLOYER-JCR	"7"	This field is checked to determine if the claimant is attached to his/her previous employer or affiliated with a union with a hiring hall. If value of the field is "7" the claimant will be included.
CPD-CLAIM-TYPE	>15	This field is checked to determine if claimant has been selected to participate in the Work Search Demonstration. If the value of the field greater than 15 the claimant will be excluded.
CPD-SG7-CTR	N/A	Used to access the segment 7 record with the earliest ISSUE-CHK-DATE.

BPH-ISSUE-CHK-DATE	Date Range	This date range is checked against the last segment seven record to extract one week of claimant first pay data. The date range is Sunday through Saturday and is stored in DATE-WEEK-START and DATE-WEEK-END.
CPD-SG4-CTR	N/A	Used to access the first segment 4 record for review.
BPE-RTW-DT	000000 or 111111	This field is checked to determine if the claimant is attached to his/her previous job. If the field is not 000000 and not set to 11/11/11, the claimant will be excluded.

DATA FIELDS TO EXTRACT: The following input fields will be loaded to the associated output fields. Any special selection criteria is listed.

Input	Output	Special Requirements
UI:CPD-SSN	SSN	JS:ES1-SSN is used to link UI and JS files.
UI:CPD-LAST	NAME-LAST	
UI:CPD-FIRST	NAME-FIRST	
UI:CPD-MIDDLE-INIT	NAME-MIDDLE-INIT	
UI:CPD-STREET	ADDR-STREET	
UI:CPD-STREET-EXT	ADDR-STREET-EXT	
UI:CPD-CITY	ADDR-CITY	
UI:CPD-STATE	ADDR-STATE	
UI:CPD-ZIP-CODE	ADDR-ZIP-CODE	
UI:CPD-STATE-COUNTY-FIPS	ORIG-SDA	
UI:CPD-TELEPHONE-NO-AC	TELEPHONE-NO-AC	
UI:CPD-TELEPHONE-NO	TELEPHONE-NO	
UI:CPD-BIRTH-DATE	BIRTH-DATE	
UI:CPD-SEX	SEX	
UI:CPD-ETHNIC-GROUP	ETHNIC-GROUP	
JS:ES1-OVET	VETERAN-DATA-IND	If JS:ES1-OVET empty Load from UI:CPD-VETERAN-DATA-IND. Loaded with value of "Y" or "N".
CC:DATE-WEEK-END	DATE-WEEK-END	(Control Card - PROFDTE)
JS:ES1-OCC	ORIG-OCC	
UI:CPD-LOCAL-OFFICE	LOCAL-OFFICE	
UI:BPE-EMPLOYER-SIC	ORIG-IND	
UI:BPE-START-DATE	START-DATE	
UI:BPE-END-DATE	END-DATE	
JS:ES1-HIGR	ORIG-HIGR	

Output file will serve as input file to the Conversion program.

CONVERSION PROGRAM: This process takes the consolidated data extracted from the JS and UI files and converts the elements to standard values before processing the data through the profiling program. This process also creates an error report that will list the invalid data encountered in the Education, Tenure, Occupation, and Industry data. Bad data will be classified into two categories; blank and invalid.

Input File Desc: Profiling Extract File
Input File 1: EMNP.ENDS.UIJS.DATA
Input Format: Sequential file format
Sort Criteria: SSN order

Input File Desc: Default Values Control Card
Input File 2: EMNP.ENDS.CTLCDS(PROFDEF)
Input Format: Partitioned Data Set file format.
Sort Criteria: Not Applicable

Input File Desc: Date Control Card
Input File 3: EMNP.ENDS.CTLCDS(PROFDTE)
Input Format: Partitioned Data Set file format.
Sort Criteria: Not Applicable

Output File Desc: Temporary Converted Extract File
Output File 1: &&TMPPRF
Output Format: Sequential file format (temporary)
Sort Criteria: SSN order

Output File Desc: Data Error Report
Output File 2: To Printer
Output Format: Report file format (See Attachment F)
Sort Criteria: SSN Order

CONVERSION CRITERIA: This section defines the input fields, the movement of input data to the output fields, and the requirements to convert the data to a new value. The section also provides validation information, the default value, and the flag settings based on the validation results.

Input	Default	Output	Selection/Conversion Criteria
ORIG-HIGR	DEF-HIGR	CONV-HIGR FLG-HIGR	Translate the Education value based on the following criteria. If blank set FLG-HIGR to B, if invalid to I. If invalid or blank load DEF-HIGR to CONV-HIGR.

High Grade	= Grade Code
00-11, GD	= 1
12	= 0
13-19, C2-C5, AD	= 2
C6-C7, BD	= 3
C8-C9, MD, PD	= 4

START-DATE END-DATE	DEF-TENURE	CONV-TENURE FLG-TENURE	Tenure is calculated to the year. If less than one year set to zero. Truncate any partial year. End Date must be equal to or later than start date. Total years between two dates cannot be greater than 60. Dates that have a START-DATE or END-DATE equal to "00/00/00", "01/01/01", "11/11/11", or "12/12/12" are invalid. If blank set FLG-TENURE to B, if invalid to I. If blank or invalid load DEF-TENURE to CONV-TENURE.
ORIG-IND	DEF-IND	CONV-IND FLG-IND	Translate the SIC (Industry) code (all 6 positions) to a 2 digit field. Use the following criteria for this translation. If blank set FLG-IND to B, if invalid to I. If blank or invalid load DEF-IND to CONV-IND. Industry Code = IND 010000-099999 = 00 100000-149999 = 01 150000-179999 = 02 200000-399999 = 03 400000-499999 = 04 500000-519999 = 05 520000-599999 = 06 600000-699999 = 07 700000-899999 = 08 910000-979999 = 09 990000-999999 = 10
ORIG-OCC	DEF-OCC	CONV-OCC FLG-OCC	Translate the DOT (Occupation) code (first 3 positions) to a 1 digit field. Use the following criteria for this translation. If blank set FLG-OCC to B, if invalid to I. If blank or invalid load DEF-OCC to CONV-OCC. Occup. =OCC 001-199 = 1 200-299 = 2 300-399 = 3 400-499 = 4 500-599 = 5 600-699 = 6 700-799 = 7 800-899 = 8 900-999 = 9

ORIG-SDA	DEF-SDA	CONV-SDA FLG-SDA	Validate that the state code, first two positions, is "24" (Maryland), then translate the county code last three positions into SDA as follows. State codes other than "24" are invalid. If blank set FLG-SDA to B, if invalid to I. If invalid or blank load DEF-SDA to CONV-SDA.

County Code	= SDA Code	
001	= 007	
003	= 011	
005	= 001	
009	= 006	
011	= 009	
013	= 012	
015	= 008	
017	= 006	
019	= 009	
021	= 005	
023	= 007	
025	= 008	
027	= 012	
029	= 009	
031	= 004	
032	= 004	
033	= 003	
035	= 009	
037	= 006	
039	= 010	
041	= 009	
043	= 007	
045	= 010	
047	= 010	
510	= 002	
997	= 999	(Default)
998	= 999	(Default)

Output file will serve as input file to the Profiling program.

PROFILING PROGRAM: This process takes the standardized data extracted from the JS and UI data files and ranks the individual claimants based on the their probability of exhausting their benefits before reemployment.

Input File Desc: Temporary Converted Extract File
Input File 1: &&TMPPRF
Input Format: Sequential file format (temporary)
Sort Criteria: SSN

Input File Desc: SDA/URATE/Industry Percent of Change File
Input File 2: EMNP.ENDS.PROFILE.LMIDATA
Input Format: Sequential Data Set
Sort Criteria: Service Delivery Area (SDA)

Input File Desc: Coefficients Control Card
Input File 3: EMNP.ENDS.CTLCDS(PROFCOEF)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Output File Desc: Final Profiled Claimant File
Output File 1: &&TMPPR2
Output Format: Sequential file format (temporary)
Sort Criteria: SSN

FORMULA CRITERIA: The formula criteria defines the calculation used to rank individual claimants. There are three elements which must have values assigned:

Exp	Is defined through statistical theory and is constant.
Bxi	Is based on claimant specific values of: Unemployment Rate and Industry % of Change within a SDA, Occupation, Tenure, and Education, and the application of the coefficients in relation to these values.
Ranking Value	Product of the equation.

Exp is defined as: 2.718281828

Bxi is determined as follows:

1. Bxi = 0 (Set to 0)

2. Add Value of Coefficient Card 1 (Add Baseline)

3. If CONV-HIGR = 0 (Add Education)
   Add Value of Coefficient Card 2
   If CONV-HIGR = 1
   Add Value of Coefficient Card 3
   If CONV-HIGR = 2
   Add Value of Coefficient Card 4
   If CONV-HIGR = 3
   Add Value of Coefficient Card 5
   If CONV-HIGR = 4
   Add Value of Coefficient Card 6

```
4.    Add Result of (CONV-TENURE multiplied by                (Add Tenure)
      Value of Coefficient Card 7)

5.    If CONV-OCC = 1                                         (Add Occupation)
            Add Value of Coefficient Card 8
      If CONV-OCC = 2
            Add Value of Coefficient Card 9
      If CONV-OCC = 3
            Add Value of Coefficient Card 10
      If CONV-OCC = 4
            Add Value of Coefficient Card 11
      If CONV-OCC = 5
            Add Value of Coefficient Card 12
      If CONV-OCC = 6
            Add Value of Coefficient Card 13
      If CONV-OCC = 7
            Add Value of Coefficient Card 14
      If CONV-OCC = 8
            Add Value of Coefficient Card 15
      If CONV-OCC = 9
            Add Value of Coefficient Card 16

6.    If CONV-SDA = SDA                          (Add Unemployment Rate)
            Add Result of (URATE multiplied by
            Value of Coefficient Card 17)
      Else next SDA parameter card and try again.


7.    If CONV-SDA = SDA
       Then If CONV-IND = 00                  (Add Industry % of Change)
                  Add Result of (IND00 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 01
                  Add Result of (IND01 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 02
                  Add Result of (IND02 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 03
                  Add Result of (IND03 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 04
                  Add Result of (IND04 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 05
                  Add Result of (IND05 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 06
                  Add Result of (IND06 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 07
                  Add Result of (IND07 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 08
                  Add Result of (IND08 multiplied by
                  Value of Coefficient Card 18)
            If CONV-IND = 09
                  Add Result of (IND09 multiplied by
                  Value of Coefficient Card 18)
```

If CONV-IND = 10
Add Result of (IND10 multiplied by
Value of Coefficient Card 18)
If CONV-IND = 99
Add Result of (IND99 multiplied by
Value of Coefficient Card 18)
Else next SDA parameter card and try again.

The full Bxi equation then looks like:

Bxi = 0 + Base Coeff + Education Coeff + (Tenure Yrs x Tenure Coeff) + Occupation Coeff + (Unemployment Rate x Unemployment Coeff) + (Industry % of Change x Industry Coeff)

Ranking Value is determined from the equation: The ranking value is then determined by applying the Exp Value and the Bxi Value into the following equation:

$$\text{Ranking Value} = \frac{\text{Exp}^{(Bxi)}}{1 + \text{Exp}^{(Bxi)}}$$

OUTPUT CRITERIA: This section defines where the results of the program will be stored. The profiling program creates one data elements RANKING and stores it in the RANKING field of the final profiling file (See output file defined above.

Output Field	Obtained From
RANKING	Result of Formula Ranking Value

Output file will serve as input to the Ranking and Management Reports.

RANKING REPORT: This report program uses the data extracted and converted from the JS and UI data files and the output from the profiling program to produce a ranked listing of individual claimants based on the their probability of exhausting their benefits before reemployment.

Input File Desc: Final Profiled Claimants File
Input File 1: &&TMPPR2
Input Format: Sequential file format
Sort Criteria: SSN

Input File Desc: Date Range Control Card
Input File 2: EMNP.ENDS.CTLCDS(PROFDTE)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Input File Desc: Local Office Control Card
Input File 3: EMNP.ENDS.CTLCDS(PROFLOFF)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Output File Desc: Ranking Report
Output File 1: To Printer
Output Format: Report file format (See Attachment F)
Sort Criteria: Local Office, Ranking

OUTPUT CRITERIA: This defines the fields to be printed on the report.

File	Field	Description
2	DATE-WEEK-BEGIN	Date for beginning of week. (Header Only)
2	DATE-WEEK-END	Date for end of week. (Header Only)
1	LOCAL-OFFICE	Local Office Number and Name (translated from Number)
1	RANKING	Ranking Value of Profiling Formula (used for storing results of calculation)
1	SSN	Social Security Number
1	NAME-LAST	Last Name (1st 15 only)
1	NAME-FIRST	First Name (1st 10 only)
1	NAME-MIDDLE-INT	Middle Initial of Name
1	TELEPHONE-NO-AC	Phone Number - Area Code
1	TELEPHONE-NO	Phone Number
1	ORIG-HIGR	Education
1	CONV-TENURE	Tenure
1	ORIG-IND	Industry (1st four only)
1	CONV-SDA	Service Delivery Area
1	ORIG-SDA	County Code (last three only)
1	ORIG-OCC	Occupation
1	VETERAN-DATA-IND	Veteran Flag

SPECIAL REQUIREMENTS:
1. Produce a header which provides the local office number and name, beginning and ending dates selected, and the system processing date.
2. Report will produce a page break when the value of the CPD-LOCAL-OFFICE field changes.
3. Report will provide remote printing routing for each local office's report information.

MANAGEMENT REPORT 1: This report program takes the standardized data extracted from the JS and UI data files and creates a statistical report of the number of claimants profiled in the state and in each local office.

Input File Desc: Final Profiled Claimants File
Input File 1: &&TMPPR2
Input Format: Sequential file format
Sort Criteria: SSN

Input File Desc: Date Range Control Card
Input File 2: EMNP.ENDS.CTLCDS(PROFDTE)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Input File Desc: Default Values Control Card
Input File 3: EMNP.ENDS.CTLCDS(PROFDEF)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Input File Desc: Local Office Control Card
Input File 4: EMNP.ENDS.CTLCDS(PROFLOFF)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Output File Desc: Number of Profiled Claimants Report
Output File 1: To Printer
Output Format: Report file format (See Attachment F)
Sort Criteria: Local Office

OUTPUT CRITERIA: This defines the fields printed on the report.

File	Field	Description
2	DATE-WEEK-BEGIN	Date for beginning of week. (Header Only)
2	DATE-WEEK-END	Date for end of week. (Header Only)
1	LOCAL-OFFICE	Local Office Name, as translated from the numerical value.
1	RANKING	Ranking Value from Profiling Formula (used for calculation)
3	DEF-BENCHMARK	Default for Reporting Benchmark (used for calculation)

SPECIAL REQUIREMENTS:
1. Tally the number of claimants, statewide and by local office.
2. Tally the number of claimants, statewide and by local office, whose ranking was equal to or higher than the benchmark.
3. Calculate the percentage of claimants, statewide and by local office, whose ranking was equal to or higher than the benchmark, as compared to the number of total claimants.

MANAGEMENT REPORT 2: This report program takes the standardized data extracted from the JS and UI data files and creates a statistical report of the number of claimants by variable which had invalid or missing data elements. This information will be reported on the state and local office levels.

Input File Desc: Final Profiled Claimants File
Input File 1: &&TMPPR2
Input Format: Sequential file format
Sort Criteria: SSN

Input File Desc: Date Range Control Card
Input File 2: EMNP.ENDS.CTLCDS(PROFDTE)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Input File Desc: Local Office Control Card
Input File 3: EMNP.ENDS.CTLCDS(PROFLOFF)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Output File Desc: Invalid Claimant Data Trends Report
Output File 1: To Printer
Output Format: Report file format (See Attachment F)
Sort Criteria: Local Office

OUTPUT CRITERIA: This defines the fields printed on the report.

File	Field	Description
2	DATE-WEEK-BEGIN	Date for beginning of week. (Header Only)
2	DATE-WEEK-END	Date for end of week. (Header Only)
1	LOCAL-OFFICE	Local Office (convert number to name for display)
1	FLG-HIGR	Education Flag - Blank or Invalid (used for tally)
1	FLG-TENURE	Tenure Flag - Blank or Invalid (used for tally)
1	FLG-OCC	Occupation Flag - Blank or Invalid (used for tally)
1	FLG-IND	Industry Flag - Blank or Invalid (used for tally)

SPECIAL REQUIREMENTS:

1. Tally the number of claimants statewide.
2. Tally the number of claimants by local office.
3. Tally the number of claimants, statewide and by local office with blank or invalid education data (FLG-HIGR)
4. Tally the number of claimants, statewide and by local office with blank or invalid tenure (FLG-TENURE)
5. Tally the number of claimants, statewide and by local office with blank or invalid occupation (FLG-OCC)
6. Tally the number of claimants, statewide and by local office with blank or invalid industry (FLG-IND)

MANAGEMENT REPORT 3: This report program takes the standardized data extracted from the JS and UI data files to create a statistical report for each variable value, which provides the number and percent of total claimants.

Input File Desc: Final Profiled Claimants File
Input File 1: &&TMPPR2
Input Format: Sequential file format
Sort Criteria: SSN

Input File Desc: Date Range Control Card
Input File 2: EMNP.ENDS.CTLCDS(PROFDTE)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Input File Desc: Default Values Control Card
Input File 3: EMNP.ENDS.CTLCDS(PROFDEF)
Input Format: Partitioned Data Set
Sort Criteria: Not Applicable

Output File Desc: Benchmark Probability Report
Output File 1: To Printer
Output Format: Report file format (See Attachment F)
Sort Criteria: Local Office

OUTPUT CRITERIA: This defines the fields printed on the report.

File	Field	Description
2	DATE-WEEK-BEGIN	Date for beginning of week. (Header Only)
2	DATE-WEEK-END	Date for end of week. (Header Only)
1	CONV-HIGR	Education
1	CONV-TENURE	Tenure
1	CONV-OCC	Occupation
1	CONV-IND	Industry
1	RANKING	Ranking Value of Prof. Formula (used in calc.)
3	DEF-BENCHMARK	Default for Reporting Benchmark (used in calc.)

SPECIAL REQUIREMENTS:

1. Tally the number of statewide claimants.
2. Tally the number of statewide claimants whose ranking was equal to or higher than the benchmark.
3. Tally the number of statewide claimants based on specific values of the following variables: Education, Tenure, Occupation, Industry.
4. Tally the number of statewide claimants based on the values of the following variables, and whose ranking was equal to or higher than the benchmark: Education, Tenure, Occupation, Industry.
5. Calculate the percentage of statewide claimants based on values of the following variables, as compared to the number of statewide claimants: Education, Tenure, Occupation, Industry.
6. Calculate the percentage of statewide claimants based on values of the following variables, as compared to the number of statewide claimants whose ranking was equal to or higher than the benchmark: Education, Tenure, Occupation, Industry.

CUMULATIVE PROCESS: This process creates a cumulative data file of all profiled claimants by appending each weeks worth of claimant data.

Input File Desc: Cumulative Profiling File
Input File 1: EMNP.ENDS.GDG.PROF94(0)
Input Format: Generation Data Group - Sequential file format
Sort Criteria: Date Week End, SSN

Input File Desc: Final Profiled Claimant File
Input File 1: &&TMPPR2
Input Format: Sequential file format (temporary)
Sort Criteria: SSN

Output File Desc: Cumulative Profiling File
Output File 1: EMNP.ENDS.GDG.PROF94(+1)
Output Format: Generation Data Group - Sequential file format
Sort Criteria: Date Week End, SSN

SPECIAL REQUIREMENTS: This section defines any special requirements for the cumulative file creation.

- The cumulative file should house one year of claimant data. The file will be named to reflect the fiscal year which is represented, and a new file name will be created on the first processing run within the new fiscal year. Example file name:
  EMNP.ENDS.GDG.PROF94 - represents profiled claimants during fiscal year 1994.

- The cumulative file will be created as a generation data group, so that errors in processing can be recovered without re-creation of the entire file. The maximum number of generations retained will be set to 10, allowing ten weeks of processing to catch errors.

- When adding the new weeks worth of claimant data to the cumulative file, the data will be appended to the bottom of the file. No sorting will be required, since the weekly file is sorted by SSN and represents the new week which would be loaded at the bottom.

ATTACHMENT A

LAYOUT FOR THE UNEMPLOYMENT INSURANCE FILE: Only the fields relevant for the profiling process are defined. Three types, those used for exclusion (not saved), those used for ranking, and those used for reporting. (EUZD.TEST.BENFTMTR.EUZ960F1, EUZ960F2, and EUZ960F3)

Segment 01:

CPD-KEY	PIC O(10)	Record Key (KEY)
CPD-SSN	PIC 9(9) COMP	Social Security Number (KEY)
CPD-SSN-SEQ	PIC 9	Record Number Per SSN (KEY)
CPD-SG4-CTR	PIC S9999 COMP	Segment 4 Counter (Exclusion)
CPD-SG7-CTR	PIC S9999 COMP	Segment 7 Counter (Exclusion)
CPD-LAST	PIC X(20)	Last Name (Report)
CPD-FIRST	PIC X(14)	First Name (Report)
CPD-MIDDLE-INIT	PIC X	Middle Initial (Report)
CPD-STREET	PIC X(35)	Street Address 1 (Info Only)
CPD-STREET-EXT	PIC X(35)	Street Address 2 (Info Only)
CPD-CITY	PIC X(20)	City (Info Only)
CPD-STATE	PIC XX	State Abbreviation (Info Only)
CPD-ZIP-CODE	PIC X(10)	Zip Code (Info Only)
CPD-STATE-COUNTY-FIPS	PIC 9(5) COMP-3	SDA/Last 3 Positions (Conversion)
CPD-RESIDENCE-CODE	PIC 9(4) COMP-3	Residence Code (Info Only)
CPD-TELEPHONE-NO-AC	PIC 9(3) COMP-3	Phone - Area Code (Report)
CPD-TELEPHONE-NO	PIC 9(7) COMP-3	Phone Number (Report)
CPD-BIRTH-DATE	PIC 9(6) COMP-3	Birth Date (Info Only) YYMMDD
CPD-SEX	PIC 9	Sex; Ina-0, Male-1, Female-2 (Info Only)
CPD-ETHNIC-GROUP	PIC 9	Race; Ina-0, White-1, Black-2, Asian-Islander-3, Indian-Alaskan-4, Hispanic-5, NW-Other-6 (Info Only)
CPD-VETERAN-DATA-IND	PIC 9	Claimant a Veteran? Ina-0, Yes-1, No-2 (Report)
CPD-EMPLOYER-JCR	PIC 9	Attached/Union Affiliated (Exclusion)
CPD-DOT-ONE	PIC X(10)	Occupation (Conversion, Profiling)
CPD-LOCAL-OFFICE	PIC 99	Local office number: Valid numbers: 1-5, 7-15, 20-27, 33-34, 36, 40, 42-43, 45, 50-56, 58-59, 91, 93-95, 97. The LO name is defined as an 88 level. (Exclusion, Reporting)
CPD-CLAIM-TYPE	PIC 99	Claim Type (Exclusion)

Segment 4: (may be multiple segment 4s for this record)

BPE-EMPLOYER-SIC	PIC 9(6)	Industry Code (Profiling, Reporting)
BPE-START-DATE	PIC 9(6) COMP-3	Tenure (Conversion, Profiling, Reporting) YYMMDD
BPE-END-DATE	PIC 9(6) COMP-3	Tenure (Conversion, Profiling, Reporting) YYMMDD
BPE-RTW-DT	PIC 9(6) COMP-3	Return to Work Date (Exclusion) YYMMDD

Segment 7: (may be multiple segment 7s for this record)

BPH-ISSUE-CHK-DATE	PIC 9(6) COMP-3	Check Issue Date (Excl) YYMMDD

ATTACHMENT B

LAYOUT FOR THE JOB SERVICES FILE: Only the fields relevant for the profiling process are defined. Three types, those used for exclusion (not saved), those used for ranking, and those used for reporting. (EMNV.ENDS.APPDATA.CASAC-AR or EMND.ENDT.VSCLST.APP)

ES1-SSN	PIC X(009)	Social Security Number (KEY)
ES1-HIGR	PIC X(002)	Education (Conversion, Profiling, Reporting)
ES1-OCC	PIC X(009)	Occupation (Conversion, Profiling, Reporting)
ES1-OVET	PIC X(001)	Veteran Indicator (Reporting); Review the values of this element to determine if veteran.

ATTACHMENT C

LAYOUT FOR THE PROFILING EXTRACT FILES: (EMNP.ENDS.UIJS.DATA, &&TMPPRF, &&TMPPR2, and, EMNP.ENDS.GDG.PROF94)

SSN	PIC 9(9)	Social Security Number (Loaded From UI:CPD-SSN)
NAME-LAST	PIC X(20)	Last Name (Loaded From UI:CPD-LAST)
NAME-FIRST	PIC X(14)	First Name (Loaded From UI:CPD-FIRST)
NAME-MIDDLE-INIT	PIC X	Middle Initial (Loaded From UI:CPD-MIDDLE-INIT)
ADDR-STREET	PIC X(35)	Street Address 1 (Loaded From UI:CPD-STREET)
ADDR-STREET-EXT	PIC X(35)	Street Address 2 (Loaded From UI:CPD-STREET-EXT)
ADDR-CITY	PIC X(20)	Address City (Loaded From UI:CPD-CITY
ADDR-STATE	PIC XX	Address State Abbreviation (Loaded From UI:CPD-STATE)
ADDR-ZIP-CODE	PIC X(10)	Address Zip Code (Loaded From UI:CPD-ZIP-CODE)
TELEPHONE-NO-AC	PIC 9(3)	Telephone Number Area Code (Loaded from UI:CPD-TELEPHONE-NO-AC)
TELEPHONE-NO	PIC 9(7)	Telephone Number (Loaded from UI:CPD-TELEPHONE-NO)
BIRTH-DATE	PIC 9(6)	Birth Date YYMMDD (Loaded from UI:CPD-BIRTH-DATE)
SEX	PIC 9	Sex (Loaded from UI:CPD-SEX) Values are: 0 - Information Not Available 1 - Male 2 - Female
ETHNIC-GROUP	PIC 9	Race (Loaded from UI:CPD-ETHNIC-GROUP) Values are: 0 - Information Not Available 1 - White 2 - Black 3 - Asian-Islander 4 - Indian-Alaskan 5 - Hispanic 6 - NW-Other
VETERAN-DATA-IND	PIC X	Veteran Indicator (Loaded From JS:ES1-OVET or UI:CPD-VETERAN-DATA-IND) Load with Y or N value depending on values in JS and UI fields.
DATE-WEEK-END	PIC 9(6)	Date Week Ended. YYMMDD format. (Loaded from PROFDTE:DATE-WEEK-END)
LOCAL-OFFICE	PIC 99	Local Office (Loaded From UI:CPD-LOCAL-OFFICE) Values are: 1-5, 7-15, 20-27, 33-34, 36, 40, 42-43, 45, 50-56, 58-59, 91, 93-95, 97.
ORIG-SDA	PIC 9(5)	Original SDA (Load From UI:CPD-STATE-COUNTY-FIPS)

CONV-SDA	PIC 9(3)	Converted SDA (Created from last three field positions of the ORIG-SDA field)
FLG-SDA	PIC X	Loaded in conversion program, based on valid check of ORIG-SDA
START-DATE	PIC 9(6)	Employment Date YYMMDD (Tenure Eval) (Loaded from UI:BPE-START-DATE)
END-DATE	PIC 9(6)	Employment Date YYMMDD (Tenure Eval) (Loaded from UI:BPE-END-DATE)
CONV-TENURE	PIC 99	Tenure, In Years (Loaded from calculation on START-DATE, END-DATE)
FLG-TENURE	PIC X	Loaded in conversion program, based on valid check of ORIG-TENURE
ORIG-HIGR	PIC X(2)	Education (High Grade) (Loaded from JS:ES1-HIGR)
CONV-HIGR	PIC X	Converted Education (Loaded from conversion of ORIG-HIGR)
FLG-HIGR	PIC X	Loaded in conversion program, based on valid check of ORIG-HIGR
ORIG-OCC	PIC X(9)	Occupation Code (DOT) (Loaded from JS:ES1-OCC. If JS:ES1-OCC empty Load from UI:CPD-DOT-ONE) CAUTION: UI:CPD-DOT-ONE IS PIC X(10)
CONV-OCC	PIC X	Occupation Code, Converted. (Loaded from conversion of ORIG-OCC)
FLG-OCC	PIC X	Loaded in conversion program, based on valid check of ORIG-OCC
ORIG-IND	PIC 9(6)	Industry (Loaded from UI:BPE-EMPLOYER-SIC)
CONV-IND	PIC XX	Converted Industry. Once converted, field represents Division. (Loaded from conversion of ORIG-IND)
FLG-IND	PIC X	Loaded in conversion program, based on valid check of ORIG-IND
RANKING	PIC S(9V99999)	Ranking Value from Profiling Formula (Loaded from the calculation performed in the Profiling Program formula)

ATTACHMENT D

LAYOUT FOR CONTROL CARDS/SEQUENTIAL FILES

DATE RANGE CONTROL CARD: (EMNP.ENDS.CTLCDS(PROFDTE))

FIELD NAME	FORMAT	DESCRIPTION
DATE-WEEK-BEGIN	PIC 9(8)	CCYYMMDD - Date for beginning of week. Always set to a Sunday date.
DATE-WEEK-END	PIC 9(8)	CCYYMMDD - Date for end of week. Always set to the date of the first Saturday after the Sunday date set in DATE-WEEK-BEGIN.
	Example Values:	**1994042419940430**

DEFAULT VALUES CONTROL CARD: (EMNP.ENDS.CTLCDS(PROFDEF))

FIELD NAME	FORMAT	DESCRIPTION
DEF-TENURE	PIC X(2)	Default Tenure Code
DEF-HIGR	PIC X	Default Education Code
DEF-OCC	PIC X	Default Occupation Code
DEF-IND	PIC X(2)	Default Industry Code
DEF-SDA	PIC X(3)	Default Service Delivery Area
DEF-BENCHMARK	PIC 9	Default for Reporting Benchmark (ranking value)
	Example Values:	**0208999996**

COEFFICIENTS CONTROL CARD: (EMNP.ENDS.CTLCDS(PROFCOEF))

FIELD	FORMAT	DESCRIPTION
COEF-NAME	X(4)	Coefficient Name
COEF-VALUE	S(9V9999)	Coefficient Value

There are 18 coefficient values defined in this control card, one coefficient per card. The coefficients values are defined as follows:

CARD	DESCRIPTION
1	Maryland Base Coefficient Name and Value
2	Coefficient Name and Value for Education Variable = 0
3	Coefficient Name and Value for Education Variable = 1
4	Coefficient Name and Value for Education Variable = 2
5	Coefficient Name and Value for Education Variable = 3
6	Coefficient Name and Value if Education Variable = 4
7	Tenure Coefficient Name and Value
8	Coefficient Name and Value for Occupation Variable = 1
9	Coefficient Name and Value for Occupation Variable = 2
10	Coefficient Name and Value for Occupation Variable = 3
11	Coefficient Name and Value for Occupation Variable = 4
12	Coefficient Name and Value for Occupation Variable = 5
13	Coefficient Name and Value for Occupation Variable = 6
14	Coefficient Name and Value for Occupation Variable = 7
15	Coefficient Name and Value for Occupation Variable = 8
16	Coefficient Name and Value for Occupation Variable = 9
17	SDA Unemployment Rate Coefficient Name and Value
18	Industry Coefficient Name and Value
	Example Values: BASE -0.7293
	HGR0 +0.0000

LOCAL OFFICE CONTROL CARD: (EMNP.ENDS.CTLCDS(PROFLOFF))

FIELD NAME	FORMAT	DESCRIPTION
LO-NUMBER	99-	Local Office Number
LO-NAME	X(15)	Local Office Name

There are 30 local office values defined in this control card, one local office per card. Example: 01-BALTIMORE
02-GLEN BURNIE

SERVICE DELIVERY AREA WITH UNEMPLOYMENT RATE AND INDUSTRY PERCENT OF CHANGE SEQUENTIAL FILE: (EMNP.ENDS.PROFILE.LMIDATA)

There are 13 SDAs: 001-012, 999; and 12 Industry Codes 00-10, 99

SDA	PIC X(3)	Service Delivery Area
URATE	PIC (99V9)	Unemployment Rate of Change
IND00-NUMB	PIC X(4)	Industry Number: 00
IND00-VALUE	PIC S(99V9999)	Industry Area 00 % of Change
IND01-NUMB	PIC X(4)	Industry Number: 01
IND01-VALUE	PIC S(99V9999)	Industry Area 01 % of Change
IND02-NUMB	PIC X(4)	Industry Number: 02
IND02-VALUE	PIC S(99V9999)	Industry Area 02 % of Change
IND03-NUMB	PIC X(4)	Industry Number: 03
IND03-VALUE	PIC S(99V9999)	Industry Area 03 % of Change
IND04-NUMB	PIC X(4)	Industry Number: 04
IND04-VALUE	PIC S(99V9999)	Industry Area 04 % of Change
IND05-NUMB	PIC X(4)	Industry Number: 05
IND05-VALUE	PIC S(99V9999)	Industry Area 05 % of Change
IND06-NUMB	PIC X(4)	Industry Number: 06
IND06-VALUE	PIC S(99V9999)	Industry Area 06 % of Change
IND07-NUMB	PIC X(4)	Industry Number: 07
IND07-VALUE	PIC S(99V9999)	Industry Area 07 % of Change
IND08-NUMB	PIC X(4)	Industry Number: 08
IND08-VALUE	PIC S(99V9999)	Industry Area 08 % of Change
IND09-NUMB	PIC X(4)	Industry Number: 09
IND09-VALUE	PIC S(99V9999)	Industry Area 09 % of Change
IND10-NUMB	PIC X(4)	Industry Number: 10
IND10-VALUE	PIC S(99V9999)	Industry Area 10 % of Change
IND99-NUMB	PIC X(4)	Industry Number: 99
IND99-VALUE	PIC S(99V9999)	Industry Area 99 Default % of Change

Example Record for Service Delivery Area 001, 002, and 999:

SDA	URAT	IND00	IND01	IND02	IND03	IND04	Cont>
001	01.6	00-00.0000	01-00.0000	02-00.3030	03-01.9777	04-03.3737	>>>>>
002	10.3	00-00.0000	01-00.0000	02-00.3031	03-01.9797	04-03.3777	>>>>>
999	11.2	00-00.0000	01-00.0000	02-00.3000	03-01.9700	03-03.3700	>>>>>

ATTACHMENT E

REPORT FORMATS

Error Report

Ranking Report

Management Report 1

Management Report 2

Management Report 3

CONVERSION ERRORS
B=BLANK / I=INVALID
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

SSN	FLG-HIGR	FLG-TENURE	FLG-OCC	FLG-IND	FLG-SDA
	I				
	I	I			
	B		I	I	
	B		I		
		I			
		I			
	B				
	B				
	B		I		
		I			
		I			
	B		I		
	B		I		
			I		
		I			
	B		I		
		I	I		I
		I			
		I			
				I	
		I			
	B		I		
	B		I		
	I	I			
				I	

CONVERSION ERRORS
B=BLANK / I=INVALID
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

SSN	FLG-HIGR	FLG-TENURE	FLG-OCC	FLG-IND	FLG-SDA
	B		I		
		I			
		I			
				I	
		I			
					I
		I			
	I				
	B		I		
		I			
		I			
	B		I		
				I	I
	B		I		
		I			
	B		I		
	B		I		
					I
		I			
				I	
		I			
		I			
		I			
				I	
	B	I	I		

PROFILE RANKING FOR
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

OCAL OFFICE: 93-LANDOVER

ROFILE RANK	SOCIAL SECURITY	LAST NAME	FIRST NAME	INIT	TELEPHONE NUMBER	EDUC	TENURE YEARS	SIC	SDA	CNTY CODE	DOT CODE	VET
).55350	[redacted]					12	00	9199	002	510	806261014	Y
).53469						11	01	5812	003	033	311472010	N
).52979						12	04	0000	003	033	203582054	N
).52027						12	02	5044	003	033	299357014	N
).51758						12	02	5411	003	033	219362010	N
).51758						12	02	5211	003	033	222687018	N
).51245							01	6513	003	033	248	N
).51094						80	10	1731	003	033	239362014	N
).50869						12	02	8221	003	033	213132010	N
0.50869						12	02	7363	003	033	203582054	N
0.50869						12	02	7376	003	033	213362010	N
0.50594						12	00	5651	003	033	279357054	N
0.49287						14	04	5399	003	033	219362010	N
0.48980						14	05	7376	003	033	213362010	N
0.48941						12	02	6531	003	033	382664010	N
0.48565						12	03	8811	003	033	309677010	N
0.48290						12	01	5812	003	033	381687014	N
0.48191						13	02	6022	003	033	201362030	N
0.47983						12	02	8361	003	033	311477030	N
0.47983						12	02	7221	003	033	355377018	N
0.47893						12	02	5541	003	033	189167018	N
0.47541						14	01	5611	003	033	221387050	N
0.47238							05	5999	011	003	000	N

PROFILING CLAIMANT COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

OCAL OFFICE	PASSING EXCLUSIONS	RANKED ABOVE BENCHMARK (.6+) COUNT	RANKED ABOVE BENCHMARK (.6+) PERCENT
STATEWIDE	1,123	130	11%
)1-BALTIMORE	9[illegible]	43	47%
)2-GLEN BURNIE	35	7	20%
)3-CUMBERLAND	36	2	5%
)4-HAGERSTOWN	1	0	0%
)5-FREDERICK	76	0	0%
)7-COLLEGE PK	27	0	0%
)8-ANNAPOLIS	36	1	2%
)9-TOWSON	125	23	18%
10-CAMBRIDGE	8	0	0%
11-CHESTERTOWN	5	2	40%
12-SALISBURY	0	0	0%
13-ELKTON	11	1	9%
14-OAKLAND	13	0	0%
15-WESTMINSTER	50	1	2%
20-WALDORF	47	0	0%
21-LEONARDTOWN	16	0	0%
22-BEL AIR	47	3	6%
23-ELLICOTT CITY	64	1	1%
24-DENTON	2	0	0%
25-EASTON	5	0	0%
26-CRISFIELD	5	0	0%
27-SNOW HILL	0	0	0%
33-PR. FREDERICK	15	0	0%

PROFILING CLAIMANT COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

LOCAL OFFICE	PASSING EXCLUSIONS	RANKED ABOVE BENCHMARK (.6+) COUNT	RANKED ABOVE BENCHMARK (.6+) PERCENT
14-GRASONVILLE	1	0	0%
16-PR. ANNE	0	0	0%
10-EASTPOINT	146	16	10%
12-OCEAN CITY	0	0	0%
43-WHEATON	148	0	0%
45-NORTHWEST	75	30	40%
93-LANDOVER	38	0	0%

INVALID / MISSING DATA COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

LOCAL OFFICE	TOTAL CLAIMANTS	INV / MISS EDUCATION	INV / MISS TENURE	INV / MISS DOT	INV / MISS SIC	INV / MISS SDA
01-BALTIMORE	91	6	13	7	2	0
02-GLEN BURNIE	35	16	10	16	0	0
03-CUMBERLAND	36	18	3	20	5	11
04-HAGERSTOWN	1	0	0	1	0	0
05-FREDERICK	76	51	1	48	2	1
07-COLLEGE PK	27	6	5	5	3	1
08-ANNAPOLIS	36	3	10	3	0	0
09-TOWSON	125	20	43	10	2	1
10-CAMBRIDGE	8	1	1	1	0	0
11-CHESTERTOWN	5	0	1	0	0	0
12-SALISBURY	0	0	0	0	0	0
13-ELKTON	11	1	1	1	4	0
14-OAKLAND	13	9	2	8	3	2
15-WESTMINSTER	50	2	17	2	1	1
20-WALDORF	47	11	14	9	4	1
21-LEONARDTOWN	16	9	6	9	1	0
22-BEL AIR	47	5	12	5	2	0
23-ELLICOTT CITY	64	2	15	2	5	0
24-DENTON	2	1	0	1	0	0
25-EASTON	5	0	2	0	0	0
26-CRISFIELD	5	4	1	4	1	0
27-SNOW HILL	0	0	0	0	0	0
33-PR. FREDERICK	15	3	9	3	0	0
34-GRASONVILLE	1	1	1	1	0	0
36-PR. ANNE	0	0	0	0	0	0

INVALID / MISSING DATA COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

LOCAL OFFICE	TOTAL CLAIMANTS	INV / MISS EDUCATION	INV / MISS TENURE	INV / MISS DOT	INV / MISS SIC	INV / MISS SDA
40-EASTPOINT	146	97	22	96	6	0
42-OCEAN CITY	0	0	0	0	0	0
43-WHEATON	148	29	34	31	15	5
45-NORTHWEST	75	8	23	9	3	0
93-LANDOVER	38	3	8	1	3	0
STATEWIDE	1,123	306	254	293	62	23

STATEWIDE PROFILING VARIABLES
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

VARIABLE	PASS EXCLUSION ALL PROFILED TOT	PERCENT	RANKED ABOVE BENCHMARK .6+ TOT	PERCENT
STATEWIDE	1,123	100%	130	100%
EDUCATION				
0	737	65%	74	56%
1	99	8%	48	36%
2	170	15%	8	6%
3	85	7%	0	0%
4	32	2%	0	0%
TENURE YEARS				
0	98	8%	14	10%
1	292	26%	27	20%
2	392	34%	32	24%
3	81	7%	11	8%
4	64	5%	5	3%
5	47	4%	4	3%
6	23	2%	3	2%
7	24	2%	4	3%
8	15	1%	0	0%
9	14	1%	3	2%
10	9	0%	3	2%
>10	64	5%	24	18%

STATEWIDE PROFILING VARIABLES
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

VARIABLE	PASS EXCLUSION ALL PROFILED TOT	PERCENT	RANKED ABOVE BENCHMARK .6+ TOT	PERCENT
STATEWIDE	1,123	100%	130	100%
DOT				
0/1	303	26%	6	4%
2	228	20%	48	36%
3	86	7%	26	20%
4	5	0%	0	0%
5	4	0%	1	0%
6	35	3%	5	3%
7	19	1%	6	4%
8	377	33%	23	17%
9	66	5%	15	11%
SIC				
00	5	0%	1	0%
01	1	0%	0	0%
02	102	9%	11	8%
03	104	9%	23	17%
04	69	6%	7	5%
05	65	5%	11	8%
06	197	17%	26	20%
07	115	10%	18	13%
08	347	30%	27	20%
09	53	4%	2	1%
10	3	0%	1	0%
99	62	5%	3	2%

CONVERSION ERRORS
B=BLANK / I=INVALID
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

SSN	FLG-HIGR	FLG-TENURE	FLG-OCC	FLG-IND	FLG-SDA
	I				
	I	I			
	B		I	I	
	B		I		
		I			
		I			
	B				
	B				
	B		I		
		I			
		I			
	B		I		
	B		I		
			I		
		I			
	B		I		
		I	I		I
		I			
		I			
				I	
		I			
	B		I		
	B		I		
	I	I			
				I	

CONVERSION ERRORS
B=BLANK / I=INVALID
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

SSN	FLG-HIGR	FLG-TENURE	FLG-OCC	FLG-IND	FLG-SDA
	B		I		
		I			
		I			
				I	
		I			
					I
		I			
	I				
	B		I		
		I			
		I			
	B		I		
				I	I
	B		I		
		I			
	B		I		
	B		I		
					I
		I			
				I	
		I			
		I			
		I			
				I	
	B	I	I		

PROFILE RANKING FOR
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

OCAL OFFICE: 93-LANDOVER

ROFILE RANK	SOCIAL SECURITY	LAST NAME	FIRST NAME	INIT	TELEPHONE NUMBER	EDUC	TENURE YEARS	SIC	SDA	CNTY CODE	DOT CODE	VET
).55350	[redacted]					12	00	9199	002	510	806261014	Y
).53469						11	01	5812	003	033	311472010	N
).52979						12	04	0000	003	033	203582054	N
).52027						12	02	5044	003	033	299357014	N
).51758						12	02	5411	003	033	219362010	N
).51758						12	02	5211	003	033	222687018	N
).51245							01	6513	003	033	248	N
).51094						80	10	1731	003	033	239362014	N
).50869						12	02	8221	003	033	213132010	N
0.50869						12	02	7363	003	033	203582054	N
0.50869						12	02	7376	003	033	213362010	N
0.50594						12	00	5651	003	033	279357054	N
0.49287						14	04	5399	003	033	219362010	N
0.48980						14	05	7376	003	033	213362010	N
0.48941						12	02	6531	003	033	382664010	N
0.48565						12	03	8811	003	033	309677010	N
0.48290						12	01	5812	003	033	381687014	N
0.48191						13	02	6022	003	033	201362030	N
0.47983						12	02	8361	003	033	311477030	N
0.47983						12	02	7221	003	033	355377018	N
0.47893						12	02	5541	003	033	189167018	N
0.47541						14	01	5611	003	033	221387050	N
0.47238							05	5999	011	003	000	N

PROFILING CLAIMANT COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

OCAL OFFICE	PASSING EXCLUSIONS	RANKED ABOVE BENCHMARK (.6+) COUNT	RANKED ABOVE BENCHMARK (.6+) PERCENT
STATEWIDE	1,123	130	11%
)1-BALTIMORE	9[illegible]	43	47%
)2-GLEN BURNIE	35	7	20%
)3-CUMBERLAND	36	2	5%
)4-HAGERSTOWN	1	0	0%
)5-FREDERICK	76	0	0%
)7-COLLEGE PK	27	0	0%
)8-ANNAPOLIS	36	1	2%
)9-TOWSON	125	23	18%
10-CAMBRIDGE	8	0	0%
11-CHESTERTOWN	5	2	40%
12-SALISBURY	0	0	0%
13-ELKTON	11	1	9%
14-OAKLAND	13	0	0%
15-WESTMINSTER	50	1	2%
20-WALDORF	47	0	0%
21-LEONARDTOWN	16	0	0%
22-BEL AIR	47	3	6%
23-ELLICOTT CITY	64	1	1%
24-DENTON	2	0	0%
25-EASTON	5	0	0%
26-CRISFIELD	5	0	0%
27-SNOW HILL	0	0	0%
33-PR. FREDERICK	15	0	0%

PROFILING CLAIMANT COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

LOCAL OFFICE	PASSING EXCLUSIONS	RANKED ABOVE BENCHMARK (.6+) COUNT	RANKED ABOVE BENCHMARK (.6+) PERCENT
14-GRASONVILLE	1	0	0%
16-PR. ANNE	0	0	0%
10-EASTPOINT	146	16	10%
12-OCEAN CITY	0	0	0%
43-WHEATON	148	0	0%
45-NORTHWEST	75	30	40%
93-LANDOVER	38	0	0%

INVALID / MISSING DATA COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

LOCAL OFFICE	TOTAL CLAIMANTS	INV / MISS EDUCATION	INV / MISS TENURE	INV / MISS DOT	INV / MISS SIC	INV / MISS SDA
01-BALTIMORE	91	6	13	7	2	0
02-GLEN BURNIE	35	16	10	16	0	0
03-CUMBERLAND	36	18	3	20	5	11
04-HAGERSTOWN	1	0	0	1	0	0
05-FREDERICK	76	51	1	48	2	1
07-COLLEGE PK	27	6	5	5	3	1
08-ANNAPOLIS	36	3	10	3	0	0
09-TOWSON	125	20	43	10	2	1
10-CAMBRIDGE	8	1	1	1	0	0
11-CHESTERTOWN	5	0	1	0	0	0
12-SALISBURY	0	0	0	0	0	0
13-ELKTON	11	1	1	1	4	0
14-OAKLAND	13	9	2	8	3	2
15-WESTMINSTER	50	2	17	2	1	1
20-WALDORF	47	11	14	9	4	1
21-LEONARDTOWN	16	9	6	9	1	0
22-BEL AIR	47	5	12	5	2	0
23-ELLICOTT CITY	64	2	15	2	5	0
24-DENTON	2	1	0	1	0	0
25-EASTON	5	0	2	0	0	0
26-CRISFIELD	5	4	1	4	1	0
27-SNOW HILL	0	0	0	0	0	0
33-PR. FREDERICK	15	3	9	3	0	0
34-GRASONVILLE	1	1	1	1	0	0
36-PR. ANNE	0	0	0	0	0	0

INVALID / MISSING DATA COUNTS
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

LOCAL OFFICE	TOTAL CLAIMANTS	INV / MISS EDUCATION	INV / MISS TENURE	INV / MISS DOT	INV / MISS SIC	INV / MISS SDA
40-EASTPOINT	146	97	22	96	6	0
42-OCEAN CITY	0	0	0	0	0	0
43-WHEATON	148	29	34	31	15	5
45-NORTHWEST	75	8	23	9	3	0
93-LANDOVER	38	3	8	1	3	0
STATEWIDE	1,123	306	254	293	62	23

STATEWIDE PROFILING VARIABLES
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 01

VARIABLE	PASS EXCLUSION ALL PROFILED TOT	PERCENT	RANKED ABOVE BENCHMARK .6+ TOT	PERCENT
STATEWIDE	1,123	100%	130	100%
EDUCATION				
0	737	65%	74	56%
1	99	8%	48	36%
2	170	15%	8	6%
3	85	7%	0	0%
4	32	2%	0	0%
TENURE YEARS				
0	98	8%	14	10%
1	292	26%	27	20%
2	392	34%	32	24%
3	81	7%	11	8%
4	64	5%	5	3%
5	47	4%	4	3%
6	23	2%	3	2%
7	24	2%	4	3%
8	15	1%	0	0%
9	14	1%	3	2%
10	9	0%	3	2%
>10	64	5%	24	18%

STATEWIDE PROFILING VARIABLES
WEEK OF 05/15/94 THRU 05/21/94

RUN DATE: 05/31/94
PAGE 02

VARIABLE	PASS EXCLUSION ALL PROFILED TOT	PERCENT	RANKED ABOVE BENCHMARK .6+ TOT	PERCENT
STATEWIDE	1,123	100%	130	100%
DOT				
0/1	303	26%	6	4%
2	228	20%	48	36%
3	86	7%	26	20%
4	5	0%	0	0%
5	4	0%	1	0%
6	35	3%	5	3%
7	19	1%	6	4%
8	377	33%	23	17%
9	66	5%	15	11%
SIC				
00	5	0%	1	0%
01	1	0%	0	0%
02	102	9%	11	8%
03	104	9%	23	17%
04	69	6%	7	5%
05	65	5%	11	8%
06	197	17%	26	20%
07	115	10%	18	13%
08	347	30%	27	20%
09	53	4%	2	1%
10	3	0%	1	0%
99	62	5%	3	2%

ATTACHMENT F

QUESTIONS/ANSWERS

EXTRACT PROGRAM QUESTIONS:

1. Do we want to create one extract program that handles both input files (JS/UI), or do we want an extract program for each input file? **One extract program which handles both inputs.**

2. Do we want to include the data conversion in the extract program? **No**

3. What data format will we use for the extract file? **Sequential**

4. What did Maryland use for the extract criteria of the week of claimant data previously? **Check Issue Date**

5. Can we use the first pay indicator to identify first pay claimants? **No, the first pay indicator is a flag that is turned on when the first payment is scheduled, and then turned off when the first payment is generated.**

6. In a memo from the Maryland office, a requirement was stated to exclude claimants, from the profiling effort, who were selected for the Maryland work search demonstration. What data element can be used to obtain this information so it can be used as an exclusion criteria in the profiling extract program? **Use the claimant type field with a value greater than 15. This identifies claimant who have been randomly selected to participate in work search demonstrations. Since these claimants have already been slated for services, profiling is not required.**

7. Are partial first pays a problem. **No, first pay represents the validity of a UI claimant for the profiling effort. Based on this, even a partial payment concludes that they are valid claimants.**

8. The DOT (Occupation) code is on both the UI and the JS data files. Which one is to be used? **The DOT code from the JS file is preferred, only use the DOT code from the UI file if blank in the JS file.**

9. How does the "01/01/01" and "11/11/11" affect the extract and exclusion process, and what does it represent? **The dates 01/01/01 and 11/11/11 are used in the Maryland office as a way to expedite check processing. Those claimants with a 11/11/11 in the return to work date (CPA-RTW-DT) should be included in profiling since this is not a valid return to work date. Handle the use of 01/01/01 or 11/11/11 in the BPE-START-DATE or BPE-END-DATE fields as invalid data.**

10. Is the sequential SSN number, part of the unemployment insurance record key, to be retained in the extract file? **No, the sequential number will be used as a reference in the extract process, but will not be retained in the extract file.**

11. How will multiple values for fields like the occupation code be handled? Only one of the values will be used in the profiling effort, the use of multiple codes is too complicated for initial implementation of this process. For fields like occupation, the code will be obtained from the JS data files, or from the characteristics section of the UI data file if not recorded in the JS data file.

12. Is veteran information printed on the profiling report, and subsequently needed in extract file? Yes, but only as a Yes/No flag. The Veteran Code is retained in both the UI and JS files, and has conflicting representations of the values stored. The inclusion of the extraction and translation of the data will be reviewed in the post implementation process.

13. Resolve the issue of different size occupation codes in the JS and UI files? The JS occupation code is 9 positions (agency standard) and the UI occupation code is 10 positions. The extra position is housed at the end of the regular occupation field. Since the profiling effort translates only the first three positions of the occupation code, the 10th position will not affect the profiling effort. No further research was performed.

14. Do we use the Union Affiliation field in the UI data structure for the exclusion process? No, if a claimant is union affiliated, this does not mean they are attached to a Union with Hiring Hall services, which is the exclusion requirement.

15. Are checks always issued on the same day? No, a check can be produced any day of the week, therefore a date range to identify a weeks worth of claimant data will be required in the extract process.

CONVERSION PROGRAM QUESTIONS:

1. Do we save the original values of the fields after the data conversion? Yes, the conversion results will be loaded to unique fields on the same extract file.

2. Are default values referenced/loaded during the conversion process or during the profiling process? The default values will be referenced and stored on the extract file during the conversion process.

3. What are the conversion requirements for the DOT (Occupation) code? Translates first 3 positions of the 9 digit field to a 1 digit code with values of 1 to 9.

5. What are the conversion requirements for the SIC (Industry) code? Translates all 6 positions of the field to a 2 digit code with values of 00 to 10. NOTE: The translation of the SIC code from a 6 position field to a 2 position field is identifying the divisions A-K. A-K values will not be used in profiling, the numbers 00-10 will be retained.

6. How is the FIPS code field translated to the Service Delivery Area. The FIPS code field is 5 positions and represents two values, the first two positions represent the state code, and the last three positions represent the county code. The state code of "24" for Maryland is validated, and the County code converted to the Service Delivery area. This conversion of county code to service delivery area is constant and therefore hard coded in the process.

PROFILING PROGRAM QUESTIONS:

1. What goes into the "Coefficient Table"? The coefficient table housed the coefficient values for the state baseline (1), education (2-6), Tenure (7), Occupation (8-16), Unemployment (17), and Industry (18). These coefficients are used in the formula for ranking the claimants.

2. How do the claimant converted values relate to the external tables of values like the Unemployment Rate and Industry Percent of Change? This has been answered in the detailed requirements definition of the Profiling Program. See that section of the document for the answer to this question.

REPORTING QUESTIONS:

1. Will the reports be written to disk and printed, or only printed? Only printed.

2. What are the requirements for the distribution of printing for this report to the local offices? The Ranking Report will be routed to each local office, who will only receive the section of the report related to that specific office.

OVERALL PROCESS QUESTIONS:

1. How many times will a claimant be profiled? 1 Time.

2. When should we create a permanent disk file in the process? Two permanent output disk files will be created during the profiling process. The extract program will create the first permanent disk file, since the input files are so large and expensive to access. The cumulative process will create the second permanent disk file, for long term retention. The cumulative process will run as the last step and after all data manipulations have occurred.

3. How should we store the weekly claimant data? The data will be stored in a cumulative file on disk, appending each week of claimant data to the existing accumulation of record.

4. How long will weekly extract file be retained? Since historical statistical analysis will probably occur, a yearly cumulative file will be created and retained on disk.

5. When should the backup be performed, stored on what medium, and how will it be done? Based on the requirement to house the data in a yearly cumulative file, a generation data group (GDG) approach will be used. The backups will be created during the cumulative process and retained on disk. The GDG limit has been set to 10 to allow ten weeks for identification and resolution of problem or errors in processing.

6. What day of the week is proposed for the processing the profiling cycle and how will this fit in a current Maryland production run? The Maryland office has two production runs, one for UI and one for JS. Profiling will be attached to the JS production run on Monday night.

7. Is there a way to generate the date field, instead of having to manually update weekly? We see the need to keep a date field control card since it allows flexibility in the frequency of the cycle. An automated update of the control card would alleviate human errors. For the first installation we will use the date control card with a manual update process and implement the automated update process after initial implementation.

APPENDIX B: SPSS CODE USED TO DEVELOP THE MARYLAND MODEL

SPSS CODE FOR MARYLAND MODEL

```
***************These lines read in the historic data file**********************

set mxwarns=100000
data list file=jun93smp
 /ssn 1-9 origloff 10-11 county 12-16 wba 17-22 (2) actamt 23-30 (2) empsic 31-
  36 higr 37-38 (a) dot3 39-41 tenure 48-50
save outfile=bigfile1.sys


*****These lines transform the data elements from the file into the***********
*****formats that will be used by the model.************************************
get file=bigfile1.sys
autorecode variables=higr
 /into educ
recode educ (1=sysmis) (2 thru 13,29=1) (14=0) (15 thru 21,24 thru 27,22=2)
 (23,28=3) (30,31=4)
formats educ (f1.0)

select if (tenure gt 0 and tenure lt 732)
compute tendec=tenure/12
compute tenyrs=trunc(tendec)
formats tenyrs (f2.0)

if (county=24510) sda=2
if (county=24003) sda=11
if (county=24005) sda=1
if (county=24021) sda=5
if (county=24039) sda=10
if (county=24045) sda=10
if (county=24047) sda=10
if (county=24027) sda=12
if (county=24013) sda=12
if (county=24031) sda=4
if (county=24033) sda=3
if (county=24009) sda=6
if (county=24017) sda=6
if (county=24037) sda=6
if (county=24015) sda=8
if (county=24025) sda=8
if (county=24011) sda=9
if (county=24019) sda=9
if (county=24029) sda=9
if (county=24035) sda=9
if (county=24041) sda=9
if (county=24001) sda=7
if (county=24023) sda=7
if (county=24043) sda=7
if (county lt 24001 or county gt 24510) sda=999
formats sda (f3.0)

if (empsic ge 010000 and empsic le 099999) ind=0
if (empsic ge 100000 and empsic le 149999) ind=1
if (empsic ge 150000 and empsic le 179999) ind=2
if (empsic ge 200000 and empsic le 399999) ind=3
if (empsic ge 400000 and empsic le 499999) ind=4
if (empsic ge 500000 and empsic le 519999) ind=5
if (empsic ge 520000 and empsic le 599999) ind=6
```

```
if (empsic ge 600000 and empsic le 699999) ind=7
if (empsic ge 700000 and empsic le 899999) ind=8
if (empsic ge 910000 and empsic le 979999) ind=9
if (empsic ge 990000 and empsic le 999999) ind=10
formats ind (f2.0)

if (dot3 ge 000 and dot3 le 199) occ=1
if (dot3 ge 200 and dot3 le 299) occ=2
if (dot3 ge 300 and dot3 le 399) occ=3
if (dot3 ge 400 and dot3 le 499) occ=4
if (dot3 ge 500 and dot3 le 599) occ=5
if (dot3 ge 600 and dot3 le 699) occ=6
if (dot3 ge 700 and dot3 le 799) occ=7
if (dot3 ge 800 and dot3 le 899) occ=8
if (dot3 ge 900 and dot3 le 999) occ=9
formats occ (f1.0)

compute wba26=(wba*26)
compute propben=(actamt/wba26)
if (propben ge 1) exst=1
if (propben lt 1) exst=0
formats exst (f1.0)

save outfile=bigfile2.sys


******These lines read in the BLS industry employment change data and match****
******with records on the existing file.**************************************

data list file='2qempch.txt'
 /sda 5-6 ind 8-9 q292 11-16 q293 18-23 pctch 25-30 (4) ratio 33-38 (4)
  wpctch 41-47 (4)
select if (not sysmis(ind))
recode sda (1=2) (2=11) (3=1) (4=5) (5=10) (6=12) (7=4) (8=3) (9=6) (10=8)
 (11=9) (12=7) (13=999)
formats sda (f3.0)
sort cases by sda ind
save outfile='2qempch.sys'

get file=bigfile2.sys
sort cases by sda ind
match files file=*
 /table='2qempch.sys'
 /by sda ind
save outfile=bigfile3.sys


****These lines read in the unemployment rates from Maryland's LMI office,*****
****exclude records containing missing or invalid data for certain elements,***
****and drop several fields not needed for estimating final equation.**********

data list file=unemp.txt
 /sda 5-6 sdatur 8-11 (1)
sort cases by sda
save outfile=unemp.sys

get file=bigfile3.sys
select if (not sysmis(ind))
select if (origloff ne 97)
select if (not sysmis(educ))
select if (sda ne 999)
select if (not sysmis(occ))
select if (not sysmis(exst))
```

```
match files file=*
 /table=unemp.sys
 /by sda
save outfile=bigfile4.sys
 /drop ssn higr county wba empsic tenure tendec propben


******These lines specify the variables used in estimating the final *********
******equation and conduct the estimation using logistic regression.**********

get file=bigfile4.sys

if (ratio ge .03) indch3=pctch
if (ratio lt .03) indch3=wpctch

logistic regression
 /variables=exst with educ tenyrs indch3 occ sdatur
 /categorical=educ occ
 /contrast (educ)=special(0 1 0 0 0
                          0 0 1 0 0
                          0 0 0 1 0
                          0 0 0 0 1)
 /contrast (occ)=special(1 0 0 0 0 0 0 0 0
                         0 1 0 0 0 0 0 0 0
                         0 0 1 0 0 0 0 0 0
                         0 0 0 1 0 0 0 0 0
                         0 0 0 0 1 0 0 0 0
                         0 0 0 0 0 1 0 0 0
                         0 0 0 0 0 0 1 0 0
                         0 0 0 0 0 0 0 0 1)
 /external
```

APPENDIX C: SPSS OUTPUT FROM ESTIMATION OF THE MARYLAND MODEL

```
01 Jun 94    SPSS Release 4.0 for Sun 4
13:21:35     SPSS for Unix -- LOCAL          Sun-4              SunOS 4.0

For SunOS 4.0            SPSS for Unix -- LOCAL            License Number 21292
This software is functional through September 30, 1999.

Try the new SPSS Release 4.0 features:

* LOGISTIC REGRESSION procedure                  * CATEGORIES Option:
* EXAMINE procedure to explore data              *   conjoint analysis
* FLIP to transpose data files                   *   correspondence analysis
* MATRIX Transformations Language                * GRAPH interface to SPSS Graph

See the new SPSS documentation for more information on these new features.

   1  0  get file=bigfile4.sys

File bigfile4.sys
   Created:  01 JUN 94 13:16:41 - 16 variables

   2  0  set length=none
   3  0  ************************************************************************
   4  0
   5  0  if (ratio ge .03) indch3=pctch
   6  0  if (ratio lt .03) indch3=wpctch
   7  0
   8  0
   9  0  logistic regression
  10  0   /variables=exst with educ tenyrs indch3 occ sdatur
  11  0   /categorical=educ occ
  12  0   /contrast (educ)=special(0 1 0 0 0
  13  0                            0 0 1 0 0
  14  0                            0 0 0 1 0
  15  0                            0 0 0 0 1)
  16  0   /contrast (occ)=special(1 0 0 0 0 0 0 0 0
  17  0                          0 1 0 0 0 0 0 0 0
  18  0                          0 0 1 0 0 0 0 0 0
  19  0                          0 0 0 1 0 0 0 0 0
  20  0                          0 0 0 0 1 0 0 0 0
  21  0                          0 0 0 0 0 1 0 0 0
  22  0                          0 0 0 0 0 0 1 0 0
  23  0                          0 0 0 0 0 0 0 0 1)
  24  0   /external
  25  0


01 Jun 94    SPSS Release 4.0 for Sun 4
13:23:29     SPSS for Unix -- LOCAL          Sun-4              SunOS 4.0

      Total number of cases:      43197 (Unweighted)
      Number of selected cases:   43197
      Number of unselected cases: 0

      Number of selected cases:                    43197
      Number rejected because of missing data:     0
      Number of cases included in the analysis:    43197
```

Dependent Variable Encoding:

Original Value	Internal Value
0	0
1	1

01 Jun 94 SPSS Release 4.0 for Sun 4
13:24:13 SPSS for Unix -- LOCAL Sun-4 SunOS 4.0

	Value	Freq	Parameter Coding (1)	(2)	(3)	(4)	(5)	(6)	(7)	(
OCC										
	1	11636	1.000	.000	.000	.000	.000	.000	.000	.0
	2	10907	.000	1.000	.000	.000	.000	.000	.000	.0
	3	4982	.000	.000	1.000	.000	.000	.000	.000	.0
	4	586	.000	.000	.000	1.000	.000	.000	.000	.0
	5	874	.000	.000	.000	.000	1.000	.000	.000	.0
	6	1933	.000	.000	.000	.000	.000	1.000	.000	.0
	7	1310	.000	.000	.000	.000	.000	.000	1.000	.0
	8	6735	.000	.000	.000	.000	.000	.000	.000	.0
	9	4234	.000	.000	.000	.000	.000	.000	.000	1.0
EDUC										
	0	21490	.000	.000	.000	.000				
**	1	8199	1.000	.000	.000	.000				
00	2	8812	.000	1.000	.000	.000				
01	3	3704	.000	.000	1.000	.000				
02	4	992	.000	.000	.000	1.000				

01 Jun 94 SPSS Release 4.0 for Sun 4
13:25:30 SPSS for Unix -- LOCAL Sun-4 SunOS 4.0

Dependent Variable.. EXST

Beginning Block Number 0. Initial Log Likelihood Function

-2 Log Likelihood 59775.204

* Constant is included in the model.

Beginning Block Number 1. Method: Enter

Variable(s) Entered on Step Number
1.. EDUC
TENYRS
INDCH3
OCC
SDATUR

Estimation terminated at iteration number 2 because
Log Likelihood decreased by less than .01 percent.

	Chi-Square	df	Significance
-2 Log Likelihood	58659.921	43181	.0000
Model Chi-Square	1115.283	15	.0000
Improvement	1115.283	15	.0000
Goodness of Fit	43217.043	43181	.0000

---------- Variables in the Equation ----------

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
EDUC			129.0053	4	.0000	.0450	
EDUC(1)	.2074	.0275	57.0149	1	.0000	.0303	1.2305
EDUC(2)	-.1455	.0268	29.4962	1	.0000	-.0214	.8646
EDUC(3)	-.1868	.0395	22.3885	1	.0000	-.0185	.8296
EDUC(4)	-.2773	.0689	16.2144	1	.0001	-.0154	.7579
TENYRS	.0233	.0019	150.0691	1	.0000	.0498	1.0235
INDCH3	-.0128	.0024	27.2357	1	.0000	-.0205	.9873
OCC			152.9225	8	.0000	.0479	
OCC(1)	.1254	.0353	12.6194	1	.0004	.0133	1.1337
OCC(2)	.2801	.0325	74.2380	1	.0000	.0348	1.3233
OCC(3)	.1646	.0383	18.4349	1	.0000	.0166	1.1790
OCC(4)	-.4042	.0886	20.8015	1	.0000	-.0177	.6675
OCC(5)	.1212	.0744	2.6535	1	.1033	.0033	1.1288
OCC(6)	-.0713	.0524	1.8518	1	.1736	.0000	.9311
OCC(7)	.1445	.0619	5.4568	1	.0195	.0076	1.1555
OCC(8)	.0087	.0399	.0470	1	.8283	.0000	1.0087
SDATUR	.0832	.0042	397.2157	1	.0000	.0813	1.0868
Constant	-.6783	.0408	276.2202	1	.0000		

01 Jun 94 SPSS Release 4.0 for Sun 4
13:32:30 SPSS for Unix -- LOCAL Sun-4 SunOS 4.0

Preceding task required 618.65 seconds CPU time; 654.93 seconds elapsed.

```
  26  0

    25 command lines read.
     0 errors detected.
     0 warnings issued.
   619 seconds CPU time.
   656 seconds elapsed time.
       End of job.
```